Essays on Private Investors' Decision-Making and Financial Advice

Inaugural-Dissertation
zur Erlangung des Doktorgrades
des Fachbereiches Wirtschaftswissenschaften
der Johann Wolfgang Goethe-Universität
Frankfurt am Main

vorgelegt von
Matthias Rumpf
aus
Braunschweig

Frankfurt am Main, 2020

To my parents
who provided everything and more
so their offspring could outgrow themselves.

Essays on Private Investors' Decision-Making and Financial Advice

Matthias Rumpf*

Abstract

For private investors it is imperative to a) understand and define their own, individual risk preferences, b) assess their financial and demographic circumstances to determine the individual risk-taking potential, and c) form and maintain a well-diversified risky portfolio. The three chapters of my thesis each match one of these three tasks. The first chapter of my thesis presents novel experimental evidence to test the existence of a potential projection bias in loss aversion, a significant determinant of investor preferences, thus matching task a). The second chapter is devoted to the determination of private investors' risk-taking potential based on their financial and socio-demographic circumstances, matching task b): In a large portfolio experiment, we examine the ability and heterogeneity of lay and professional advisors in matching investor demographics, such as age and income, with risky asset portfolio shares. The third and final chapter addresses the question on how to reach and maintain an efficient risky portfolio, therefore matching task c): It analyzes a decision support system for private investors that allows its users to simulate any arbitrary set of securities, and by reporting aggregated expected return and risk, to optimize their current portfolio.

*. I thank the Faculty of Economics and Business and the Graduate School of Economics, Finance and Management (GSEFM) at Goethe University Frankfurt for ten years of excellent education and an ideal research environment; Prof. Dr. Andreas Hackethal and Prof. Dr. Thomas Otter for many years of valuable advice and generous funding; Andreas Hackethal for pointing me towards a career in Data Science and Thomas for showing me the way of the Bayesians; Konstantin, the best office and climbing mate ever, especially for taking over EFL duties, including the spring conference; Christine and Steffen for our joint research endeavors; Marina, Jenny, Benni, Philipp, Sebastian, Gregor, Thomas, and Dominique for making work at the Chair for Personal Finance awesome; Anni for being the devine soul of our chair; Kristine Rach-Emma, Beate Geschwandtner, Carina Schneider, Heidrun Schneider, and Selma Ullrich for outstanding administrative and Michael Rohrdrommel for superior IT support. Special thanks goes to Prof. Michael Binder for the admission to the MSQE/PhD Program. Above all, I thank Giuli for more than ten years' love and support.

Contents

Introduction and summary

This thesis involves household investment behavior, challenges in forming efficient risky portfolios, related mistakes, and possible remedies in the form of financial advice and digital decision support systems.

Portfolio allocation is a notoriously difficult problem for private investors: Their optimal portfolio choice depends on a myriad of uncertain, time-varying financial and socio-demographic factors. It is especially challenging because of the underlying multi-period optimization problem involving not only today's, but also future consumption, savings, and risk-taking decisions over the entire life span of the investor (see e.g. Cocco et al. (2005)). Among other factors, portfolio choice depends on human capital and its age-related life-cycle dynamics (Bodie et al. (1992), Ameriks and Zeldes (2004)), income, other background risks (Heaton and Lucas (2000)), and consumption habits (Gomes and Michaelides (2003), Polkovnichenko (2007)).

Individual investors are likely not able to cope with such complexity. Actually, even the fundamental concept of risk seems too abstract for regular people, given that they benefit from, and in many cases might even require, assistance in estimating and understanding the risk of potential investments and even their own preferences for risk (see e.g. Kaufmann and Weber (2013), Bradbury et al. (2015)). An example of uncertain preferences is loss aversion (Kahneman and Tversky (1979)), which causes return variations to affect people more severely in the loss domain than in the gain domain. Recent research questions whether loss aversion is indeed a preference or a bias caused by incorrect expectations about the hedonic impact of future losses, i.e., a projection bias (Merkle (2020)).

In addition to struggling with the riskiness of investments, when facing actual investment decisions, investors often suffer from a multitude of behavioral biases causing under-diversification such as the home bias (French and Poterba (1991)) and lottery stock preferences (Kumar (2009)). Furthermore, given that some investors are not even able to aggregate their total portfolio's performance over multiple assets (Glaser and Weber (2007)), it is not surprising that investors are narrow framing (Tversky and Kahneman (1985)), i.e., they make isolated myopic investment decisions, another source of portfolio inefficiency and under-diversification.

Financial advice, probably the most obvious solution to help private investors, has thus far been a disappointing remedy. This may be the result of catering to biases and deteriorating even diversified portfolios (Mullainathan et al. (2012)) or not being adapted or followed properly (Bhattacharya et al. (2012)). Furthermore, advisors who apply one-size-fits-all heuristics (Foerster et al. (2017)) do not sufficiently incorporate the heterogeneity in financial and socio-demographic factors among individual investors.

It is hence obvious to look at three fundamental problems of private investors. For investors it is imperative to a) understand and define their own, individual risk preferences, b) assess their financial and demographic circumstances to determine the individual risk-taking potential, and c) form and maintain a well-diversified risky portfolio. The first two tasks jointly determine the optimal risky asset portfolio share, and the third task ensures efficiency of this risky portfolio. The three chapters of my thesis each match one of the three tasks.

The first chapter of this thesis presents novel experimental evidence to test the existence of a potential projection bias in loss aversion, a significant determinant of investor preferences, thus matching task a). Only recently, has research questioned whether people really suffer from loss aversion as severely as they anticipate. Merkle (2020) and Kermer et al. (2006) claim that loss aversion is a projection bias in the prediction of satisfaction, as a measure of utility, with an outcome realization in the future. Kermer et al. (2006) elicited subjective ratings on satisfaction with outcomes in a computer-simulated card-gambling experiment, and Merkle (2020) elicited investor satisfaction in a panel survey among clients of a large UK bank. Both articles found that people anticipate feeling emotional distress from suffering losses, but realized losses turn out less painful than expected. This implies that there is a significant difference between the *anticipated satisfaction* and the *realized satisfaction.* A possible explanation is that people tend to overestimate the hedonic impact of losses and underestimate their ability to cope with them. Consequently, the predicted hedonic impact attributed to losses is too large, and when people actually do experience a loss, the realized impact on utility is smaller than predicted.

If loss aversion is more present in anticipation than after realization of returns, this can have a substantial effect on how investors react to previous investment outcomes and on their risk taking behavior in general. Furthermore, as Kermer et al. (2006) point out, loss aversion would be a wealth-maximizing error. Investors suffering from a projection bias would invest more conservatively than necessary to match their preferences, thereby foregoing higher premia on the returns of riskier financial assets. However, loss aversion is not a behavioral bias as long as it reflects true preferences over (realized) outcomes.

In this chapter, I analyze the potential projection bias in satisfaction with investment outcomes, i.e., differences in anticipated and realized satisfaction, and its impact on future investment behavior. While previous work has estimated loss aversion coefficients on aggregate data separately for anticipated and realized returns across individuals, this study reports differences in satisfaction with the same outcome when evaluated ex-ante (anticipated) versus ex-post (realized) in the same individual. This granular view allows for a more detailed investigation on the root of the projection bias and its implications for investment behavior. A series of three experiments were conducted to estimate and test the existence of a projection bias in loss aversion and a projection bias in the underlying satisfaction with returns. I used a novel experimental design, which allowed me to elicit participants' subjective ratings of the same return outcome, both before (in *anticipation*) and after actually experiencing it (in *realization*). The experimental data does not confirm a projection bias in loss aversion found in previous research. Merkle (2020) report anticipated loss aversion estimates close to 2 and realized loss aversion close to 1. My loss aversion estimates on both anticipated and realized returns vary between 0.904 and 1.145 over three experiments. I find neither convincing evidence for loss aversion estimated on satisfaction with investment returns, nor for a projection bias in loss aversion, i.e., coefficients that are larger for anticipated compared to realized returns.

If a projection bias in loss aversion exists, it would originate from a gap between realized and anticipated satisfaction for investment outcomes in the loss domain. The gap could be justified, among other explanations, by peoples' tendency to underestimate their coping mechanisms (Loewenstein et al. (2003)). The results of my experiments show that average satisfaction levels for anticipated and realized returns move closely together over the full range of returns. A significant but small projection bias in satisfaction for the loss domain is identified only in one of the three experiments. In Experiment II, the average difference between realized and anticipated satisfaction in the loss domain (for returns smaller than or equal to 0%) is equal to 3.6 and significant at the 1% level. Notably, a projection bias in satisfaction and hence also in loss aversion is absent when eliciting the return satisfaction independent from the investment decision, which was achieved in Experiment III. However, this finding could be attributed to the simple risky asset return distribution with only six possible returns of equal probability.

I conclude that the projection biases in loss aversion (Merkle (2020)) and in satisfaction with outcomes (Kermer et al. (2006)) have to be tested further. The loss aversion estimates close to 1 even for anticipated satisfaction are puzzling. A possible explanation is that the investment task was not sufficiently realistic. The web-based experimental software only allowed for a virtual endowment, and actual payouts were only conducted based on a lottery after the end of the experiments. Hypothetical returns on a virtual endowment might have been too abstract to bring forth the emotions at play during real life investment experiences. Nevertheless, I propose to gather more evidence before demoting loss aversion from a preference to a behavioral bias (see Merkle (2020)).

The second chapter is devoted to determining private investors' risk-taking potential based on their financial and socio-demographic circumstances, hence matching task b). It is joint work conducted with Thomas Otter (Goethe University Frankfurt, Germany). In a large portfolio experiment, we examined the ability and heterogeneity of lay and professional advisors in matching investor demographics, such as age and income, with risky asset portfolio shares.

The leading questions are: How do professional financial advisors match client characteristics with optimal risky asset shares? How heterogeneous are the portfolio allocation rules of professionals? How is the heterogeneity in recommendations related to observed advisor characteristics? And finally, how do allocation rules of professional and lay advisors differ?

Almost everyone[1] expects financial advisors to tailor recommendation to their individual situation and personal needs. While financial advice has been found to improve diversification in client portfolios (see e.g. Bluethgen et al. (2008), Hackethal et al. (2012)), we might expect mis-aligned incentives to induce advisors to sell more risky and costly high-commission products (Inderst and

1. 97% of participants in a survey among 1,026 German adults (Net Fonds and Toluna (2015))

Ottaviani (2009)) regardless of our financial and socio-demographic situation. And indeed, in a large audit study, advisors did not mitigate obvious behavioral biases of their clients, but catered to detrimental mistakes such as return chasing and tilted even passive low cost portfolios towards inefficient, actively managed high-cost products (Mullainathan et al. (2012)). Nevertheless, recent research indicates that financial advisors are not necessarily intentionally abusing their clients' trust; rather, they might be steered by "misguided beliefs." Using a large brokerage data set with more than 10,000 financial advisors and 800,000 clients, Linnainmaa et al. (2018) found that clients mirror the investment mistakes of their advisors who deteriorate their own portfolio returns by frequent trading, return chasing, under-diversification and investing in expensive actively managed funds. In an earlier study on the same data, Foerster et al. (2017) reported that receiving advice from a bottom-decile compared to a top-decile advisor results in 1.2% lower gross returns. They further state that the advisor's fixed effects are significantly related to the advisors' characteristics, age, and risk tolerance.

Entering a clinic as a patient to receive medical attention, we certainly hope that our doctor takes at least some of our personal characteristics into account when prescribing a treatment. Likewise, as investors, we hope that financial advice goes beyond recommending a replication of the advisor's personal portfolio (Foerster et al. (2017)). However, the matching of client characteristics to optimal portfolio allocations is an important aspect of financial advice that has to date received less attention. Filling this gap, we add to the research on financial advice as well as the large body of literature on the socio-demographic determinants of portfolio choice.[2] In this essay, we used data on portfolio allocation recommendations collected in a web-based experiment among 424 independent financial advisors, which we call "professional advisors", and 450 regular people recruited "from the street" referred to as "lay advisors". Each participant stated optimal risky asset portfolio shares for five virtual clients described by financial and socio-demographic household characteristics. For the recommendations we elicited the participants' return expectations, risk preferences, demographic variables, and information about their own portfolios. Our experimental set-up overcame the endogeneity problems encountered in administrative data sets. Household characteristics of hypothetical clients were randomly generated and assigned to advisors. Furthermore, client profiles show little or no correlation across characteristics.

Our results and contributions are threefold. First, we identified the determinants of the participants' portfolio allocation rules and tested normative portfolio predictions. Our estimates showed that the clients' risk tolerance, age, and income were the most important factors in the advisors' portfolio rules. Surprisingly, wealth, whether financial or real estate, only played a secondary role judging by the effect size. Second, we found that in addition to the advisors' own risky allocation, as emphasized by Linnainmaa et al. (2018), age, long-term return expectations, and risk tolerance directly increased the recommended risky share by roughly 6% to 9% for an average client after an increase close to one standard deviation in the corresponding advisor char-

2. See for example Mankiw and Zeldes (1991) and Haliassos and Bertaut (1995) on stock market participation; Bodie and Dwight B. Crane (1997), Heaton and Lucas (2000), and Bertaut and Starr-McCluer (2000) on portfolio allocation.

acteristics. When incorporating the information on household demographics, advisors seemed to disagree primarily on how to map financial wealth into risky asset shares. Advisors' age, return expectations, and risk tolerance identified advisor groups with heterogeneous opinions in how client characteristics should be mapped into the optimal allocations. Concerning differences in recommendation heterogeneity across households, we found that very young, risk tolerant, or wealthy clients seemed to receive congruent advice to hold large risky asset shares, while poor, very risk averse and old clients were more likely to face larger heterogeneity in recommendations. Comparing professional and lay advisors, we found that professionals add value by incorporating client characteristics stronger and more reliably into their portfolio recommendations. Testing for differences in "one-size-fits-all" heuristics (a term coined by Linnainmaa et al. (2018)) between professionals and lay advisors, we found no convincing evidence that professionals incorporate their beliefs or own investment strategies more strongly into recommendations than lay advisors. Third and finally, we demonstrated the application of a parsimonious Bayesian variable selection method that allows for the investigation of heterogeneous effects on a large set of interactions using a large N small T data set. This approach is certainly applicable to an array of different topics within and outside of economics, such as the analysis of heterogeneity in treatments across medical practitioners.

The first two chapters are both related to the problem of determining the optimal risky asset portfolio share, either based on preferences or on financial and socio-demographic circumstances. The third and final chapter addresses the question of how to reach and maintain an efficient risky portfolio, therefore matching task c). It analyzes a decision support system for private investors that allows its users to simulate any arbitrary portfolio to receive information on the aggregate portfolio level in expected return and risk, enabling users to optimize portfolios formed by an array of individual investment decisions. This chapter was worked on jointly with Steffen Meyer (University of Southern Denmark).

The benefits of portfolio diversification have been well-known in finance since the introduction of Markowitz' portfolio theory (Markowitz (1952)). Nevertheless, there is abundant evidence that individual investors fail to diversify properly and thus hold inefficient portfolios. For example, investors often prefer domestic over international stocks (French and Poterba (1991)), hold concentrated lottery stock portfolios (Kumar (2009)), and even ETF investors tend to under-diversify by choosing niche products (Bhattacharya et al. (2017)). Thus far, financial advice as a possible remedy has been a disappointing solution. For example because some cater to biases deteriorating even diversified portfolios (Mullainathan et al. (2012)), and others are not accepted or followed thoroughly (Bhattacharya et al. (2012)).

Digital decision support could fill this gap, given that a large share of private investors could benefit from tools that aid decision-making and trading. Since the population of self-directed investors is steadily growing, computer tools that aid decision-making are a promising innovation contributing to the recent advance to "restore rationality" in consumer finance (Campbell (2016)), which still requires empirical evidence from the field. However, these tools do currently hardly exist. According to the FCA, in 2015 only 15% of UK advisory firms significantly used tools that “aid decision making and transacting” of their customers, while 46% did not use any. In computer science such tools are known as decision support systems (DSS). In consumer finance, research on DSS is scarce. For example, Looney and Hardin (2009) use an experimental website to test the design of decision support for retirement accounts to increase investor risk taking. In contrast, our research is aimed at active investors who hold risky, yet often inefficient portfolios.

We have conducted a field study together with a German online bank that has launched a portfolio simulation tool for its active, self-directed investors. The simulator calculates the expected aggregate risk and return of any arbitrary portfolio. Up to three simulated portfolios can be composed and displayed for comparison next to the current portfolio. We tracked 44,010 clients of whom 707 used the tool and traded simulated positions (called Sim-Traders); 2,521 used the tool, i.e. simulated at least three portfolios on a single day, but did not trade simulated position (Sim-Users) and 34,067 clients who did not use the simulation tool at all (Non-Users). Our data includes portfolio holdings and transactional trading data from June 2013 until October 2015. Given that the tool was introduced in June 2014, we observed 12 months prior to the introduction of the tool and 16 months post-tool introduction.

It is important to understand the distinction between two sets of portfolio efficiency metrics. First, we analyzed the short-term simulator metrics visible to users of the simulation tool. Simulator metrics are the aggregated monthly expected portfolio return and value-at-risk from which we additionally derive the simulator’s portfolio Sharpe ratio. The metrics use short-term expectations since they are based on only six months of historical data. Second, we used the long-run relative Sharpe ratio loss (RSRL) adapted from earlier research on diversification (Calvet et al. (2007) and Gaudecker (2015)) as our main objective measure of long term portfolio efficiency. We estimated the RSRL based on ex-ante expected returns using a multi-asset benchmark and the Capital Asset Pricing Model (CAPM). The RSRL dominates conventional measures of concentration such as the Herfindahl-Hirschman-Index (HHI) to evaluate portfolio diversification, by taking into account correlations in security returns and directly comparing an investor’s aggregate portfolio to the mean-variance efficient benchmark. Throughout the paper, we grouped investors into quartiles according to their average pre-tool introduction RSRL in order to investigate heterogeneous treatment effects across levels of ex-ante portfolio efficiency.

To assess whether investors actively optimized portfolio efficiency, when receiving the aggregate information from the simulator, we compared simulated portfolios in terms of the simulation tool metrics with the actual holdings, i.e., the starting portfolios. We determined that investors unambiguously searched for higher expected returns. Accepting higher risk, investors improved Sharpe ratios in terms of simulation tool metrics with the exception of investors in the fourth (least efficient) quartile of the RSRL distribution.

Primarily, we questioned whether following the simulation tool’s information indeed improved objective portfolio efficiency. We used ex-ante efficiency measures, the RSRL, and the underlying expected returns and portfolio standard deviation, and analyzed differences between (traded) simulations and the actual starting portfolios. A small deterioration in efficiency emerged for Q1 and Q2 investors but we found considerable efficiency gains for the less efficient investors in Q3 and Q4. Notably, starting from very high risk levels above 40% standard deviation, Q4 investors did not achieve efficiency gains by diversifying and reducing portfolio risk, but rather by achieving higher expected returns while accepting (or searching for) higher risk levels. While Q1 and Q2 investors improved efficiency in terms of the simulation metrics' Sharpe ratio, they deteriorated efficiency in terms of objective, long run ex-ante Sharpe ratio and RSRL.

We conclude that the portfolio optimization tool provides a sandbox to simulate alternatives and choose the most preferable option at zero cost. In case of the clients holding the least efficient portfolios, this results in even stronger risk taking, while potentially ignoring salient efficiency gains. Portfolio optimization using the short-term simulator metrics, therefore, does not help to mitigate biases on a long-term investment horizon. A simple scatter plot of changes in short-term versus long-term metrics of simulations over starting portfolios reveals that there is no reliable relation between the two metrics. The information provided by the simulation tool makes the resolution of under-diversification infeasible. However, given the strong preference for higher returns among all investor groups, we doubt that long-term metrics would achieve this goal. Nudging investors toward more efficient portfolios requires a more sophisticated design for a decision support tool.

Portfolio choice is an overwhelmingly complex problem and more work has to be done to restore rationality in private portfolio choice. In the first chapter, investment decisions are only affected by previous investment success if participants have the information directly in front of them. If future research confirms a projection bias in loss aversion, it remains an open question as to how to make investors aware of their bias and how to effectively encourage a change in their investment behavior. In the second chapter, we learn that financial advisors are valuable, because they are potentially able to help lay investors to better incorporate their financial and socio-demographic background into the optimal portfolio allocation. However, the strong heterogeneity in recommendations for the most difficult cases reveals the limits of human decision-making. Decision support systems are thus a possible remedy to assist lay investors and professionals alike. The success of such tools lies in the details. As shown in the third chapter, decision support systems must be carefully designed and frequently reassessed to ensure that decisions based on the system align with the users' long-term benefits. Nevertheless, such interactive tools are a promising way to help private investors identify their own preferences, to reveal biases and inefficiencies, make them salient, and finally lead users to better long-term decision-making.

Bibliography

Ameriks, John, and Stephen P. Zeldes. 2004. "How Do Household Portfolio Shares Vary with Age?" *Working Paper.*

Bertaut, Carol C., and Martha Starr-McCluer. 2000. "Household Portfolios in the United States". *SSRN Electronic Journal.*

Bhattacharya, Utpal, Andreas Hackethal, Simon Kaesler, Benjamin Loos, and Steffen Meyer. 2012. "Is Unbiased Financial Advice to Retail Investors Sufficient? Answers from a Large Field Study". *Review of Financial Studies* 25 (4): 975–1032.

Bhattacharya, Utpal, Benjamin Loos, Steffen Meyer, and Andreas Hackethal. 2017. "Abusing ETFs". *Review of Finance* 21 (3): 1217–1250.

Bluethgen, Ralph, Andreas Gintschel, Andreas Hackethal, and Armin Mueller. 2008. "Financial Advice and Individual Investors' Portfolios". *Working Paper.*

Bodie, Zvi, and Dwight B. Crane. 1997. "Personal Investing: Advice, Theory, and Evidence from a Survey of TIAA-CREF Participants". *Working Paper.*

Bodie, Zvi, Robert C. Merton, and William F. Samuelson. 1992. "Labor supply flexibility and portfolio choice in a life cycle model". *Journal of Economic Dynamics and Control* 16 (3-4): 427–449.

Bradbury, M. A. S., T. Hens, and S. Zeisberger. 2015. "Improving Investment Decisions with Simulated Experience". *Review of Finance* 19 (3): 1019–1052.

Calvet, Laurent E., John Y. Campbell, and Paolo Sodini. 2007. "Down or Out: Assessing the Welfare Costs of Household Investment Mistakes". *Journal of Political Economy* 115 (5): 707–747.

Campbell, John Y. 2016. "Restoring Rational Choice: The Challenge of Consumer Financial Regulation". *American Economic Review* 106 (5): 1–30.

Cocco, João F., Francisco J. Gomes, and Pascal J. Maenhout. 2005. "Consumption and Portfolio Choice over the Life Cycle". *Review of Financial Studies* 18 (2): 491–533.

Foerster, S., J. T. Linnainmaa, B. T. Melzer, and A. Previtero. 2017. "Retail Financial Advice: Does One Size Fit All?" *The Journal of Finance* 72 (4): 1441–1482.

French, Kenneth R., and James M. Poterba. 1991. "Investor Diversification and International Equity Markets". *American Economic Review* 81 (2): 222–226.

Gaudecker, Hans-Martin von. 2015. "How Does Household Portfolio Diversification Vary with Financial Literacy and Financial Advice? Internet Appendix". *The Journal of Finance* 70 (2): 489–507.

Glaser, Markus, and Martin Weber. 2007. "Why inexperienced investors do not learn: They do not know their past portfolio performance". *Finance Research Letters* 4 (4): 203–216.

Gomes, Francisco, and Alexander Michaelides. 2003. "Portfolio Choice With Internal Habit Formation: A Life-Cycle Model With Uninsurable Labor Income Risk". *Review of Economic Dynamics* 6 (4): 729–766.

Hackethal, Andreas, Michael Haliassos, and Tullio Jappelli. 2012. "Financial advisors: A case of babysitters?" *Journal of Banking & Finance* 36 (2): 509–524.

Haliassos, Michael, and Carol C. Bertaut. 1995. "Why Do So Few Hold Stocks?" *Economic Journal* 105 (432): 1110–1129.

Heaton, John, and Deborah Lucas. 2000. "Portfolio Choice and Asset Prices: The Importance of Entrepreneurial Risk". *The Journal of Finance* 55 (3): 1163–1198.

Inderst, Roman, and Marco Ottaviani. 2009. "Misselling through Agents". *American Economic Review* 99 (3): 883–908.

Kahneman, Daniel, and Amos Tversky. 1979. "Prospect Theory: An Analysis of Decision under Risk". *Econometrica* 47 (2): 263.

Kaufmann, Christine, and Martin Weber. 2013. "Sometimes less is more – The influence of information aggregation on investment decisions". *Journal of Economic Behavior & Organization* 95:20–33.

Kermer, Deborah A., Erin Driver-Linn, Timothy D. Wilson, and Daniel T. Gilbert. 2006. "Loss aversion is an affective forecasting error". *Psychological science* 17 (8): 649–653.

Kumar, Alok. 2009. "Who Gambles in the Stock Market?" *The Journal of Finance* 64 (4): 1889–1933.

Linnainmaa, Juhani, Brian Melzer, and Alessandro Previtero. 2018. "The Misguided Beliefs of Financial Advisors". *Journal of Finance (forthcoming).*

Loewenstein, George, Ted O'Donoghue, and Matthew Rabin. 2003. "Projection Bias in Predicting Future Utility". *The Quarterly Journal of Economics* 118 (4): 1209–1248.

Looney, Clayton Arlen, and Andrew M. Hardin. 2009. "Decision Support for Retirement Portfolio Management: Overcoming Myopic Loss Aversion via Technology Design". *Management Science* 55 (10): 1688–1703.

Mankiw, N.Gregory, and Stephen P. Zeldes. 1991. “The consumption of stockholders and non-stockholders”. *Journal of Financial Economics* 29 (1): 97–112.

Markowitz, Harry. 1952. “Portfolio Selection”. *The Journal of Finance* 7 (1): 77–91.

Merkle, Christoph. 2020. “Financial Loss Aversion Illusion”. *Review of Finance* 24 (2): 381–413.

Mullainathan, Sendhil, Markus Noeth, and Antoinette Schoar. 2012. “The Market for Financial Advice: An Audit Study”. *NBER Working Paper,* number 17929.

Net Fonds and Toluna. 2015. “Servicemonitor Finanzen: Industry Study”. Visited on September 18, 2018. `https://service.netfonds.de/marketing/pdf/Studie_Servicemonitor_B2C_2015.pdf`.

Polkovnichenko, Valery. 2007. “Life-Cycle Portfolio Choice with Additive Habit Formation Preferences and Uninsurable Labor Income Risk”. *The Review of Financial Studies* 20 (1): p 83–124.

Tversky, Amos, and Daniel Kahneman. 1985. “The Framing of Decisions and the Psychology of Choice”. In *Behavioral Decision Making,* edited by George Wright, 82:25–41. Boston: Springer US.

Chapter I.

Is loss aversion a projection bias? New experimental evidence on investors' return satisfaction

Is loss aversion a projection bias? New experimental evidence on investors' return satisfaction

Matthias Rumpf‡

Abstract

In this paper, I analyze a potential projection bias in loss aversion and the underlying satisfaction with investment returns. I conduct a series of three experiments to estimate and test the existence of the projection bias. In all three experiments I let participants decide how to allocate an endowment of $1,000 between a risky and a risk free asset. I use an experimental design which allows me to elicit participants' subjective ratings of the same return outcome, both before (in *anticipation*) and after actually experiencing the outcome (in *realization*). I cannot confirm a projection bias in loss aversion found in previous research. I find convincing evidence neither for loss aversion estimated on satisfaction with investment returns nor for a projection bias in loss aversion, i.e. coefficients that are larger for anticipated compared to realized returns. Anticipated and realized satisfaction do not consistently divert from each other in the loss domain. My findings might be attributed to a higher level of involvement in the investment task and a more instantaneous measurement of the realized satisfaction with returns compared to previous research.

‡. Goethe University Frankfurt

Introduction

Loss aversion describes the common preference among people to avoid losses. 'Looming larger than gains' on people's well-being (Kahneman and Tversky (1979)), the hedonic impact of a loss is about twice as big as an equally sized gain (Kahneman and Snell (1992)). For example, most people accept a coin flipping gamble only if the winning outcome is substantially larger than the losing outcome, in other words only if the gamble has a positive expected outcome. On the one hand, loss averse individuals only choose options that offer an acceptable compensation for possible losses, and on the other hand, they even make an effort or take risks to avoid possible losses (see e.g. Imas (2016)).

Only recently, research has questioned whether people really suffer from loss aversion as strongly as they anticipate. Merkle (2020) and Kermer et al. (2006) claim that loss aversion is a projection bias in the prediction of the "realized satisfaction" that people expect to experience for an outcome in the future. If loss aversion is more present in anticipation than after realization of returns, this can have substantial effects on how investors react to previous investment outcomes. Furthermore, as Kermer et al. (2006) point out, loss aversion would be a wealth-maximizing error. However, loss aversion is not a behavioral bias as long as it reflects true preferences over (realized) outcomes.

In this paper, I analyze a potential projection bias in satisfaction with investment outcomes, i.e. differences in anticipated and realized satisfaction, and their impact on future investment behavior. I conduct a series of three experiments to estimate and test the existence of a projection bias in loss aversion and a projection bias in the underlying satisfaction with returns. I use a novel experimental design which allows me to elicit participants' subjective ratings of the same return outcome, both before (in *anticipation*) and after actually experiencing it (in *realization*). To understand the difference between prediction and experience of satisfaction, I distinguish two types of utility, i.e. measures of satisfaction with an outcome: "Anticipated satisfaction" is the hedonic value that people predict they will receive when a certain event occurs and "realized satisfaction" is the hedonic value received immediately after the event took place.[1]

In all three experiments I let participants decide how to allocate an endowment of US$1,000 (€1,000 for experiments I and III) between a risky and a risk free asset. Participants were educated about the risky asset's return distribution, using different *information treatments.* After the initial allocation, participants were asked to give subjective ratings, called *anticipated satisfaction* on returns from risky assets' distribution. At the end of the experiment, a market simulation drew one random return out of the previously described return distribution. Participants were informed about their investment return and then asked about their satisfaction with the realized return, the *realized satisfaction.*

1. To avoid name confusion with "expected returns" or "experience sampling", I adapt but slightly alter the terminology of Merkle (2020) by using "anticipated satisfaction", which corresponds to Merkle's term "expected happiness", and "realized satisfaction" corresponding to "experienced happiness".

Economics students from Radboud University Nijmegen participated in experiments I and III, and I acquired participants via Amazon's Mechanical Turk for experiment II.

I present a simple risky asset with either two (in experiments I and II) or six possible returns (experiment III) and a complex risky asset with a continuous distribution (experiment II). In experiments I and II the additional information treatment varies between a description of and experience sampling from the risky asset's distribution. In experiment III all participants receive the description of the risky asset's distribution. In my data, I cannot confirm a projection bias in loss aversion found in previous research. Merkle (2020) reports anticipated loss aversion estimates close to 2 and realized loss aversion close to 1. My loss aversion estimates on both anticipated and realized returns vary between 0.904 and 1.145 over three experiments. I find neither convincing evidence for loss aversion estimated on satisfaction with investment returns, nor for a projection bias in loss aversion, i.e. coefficients that are larger for anticipated compared to realized returns.

If a projection bias in loss aversion exists, it would originate from a gap between realized and anticipated satisfaction for investment outcomes in the loss domain. This gap would be justified, among other explanations, by people's tendency to underestimate their coping mechanisms (Loewenstein et al. (2003)). The results of my experiments show that average satisfaction levels for anticipated and realized returns are very close over the full range of returns. A significant but small projection bias in satisfaction for the loss domain is identified only in experiment II, which asked for satisfaction on returns of a complex risky asset with a continuous return distribution among participants recruited on Amazon's Mechanic Turk platform. The average difference between realized and anticipated satisfaction in the loss domain (for returns smaller than or equal to 0%) is equal to 3.6 and significant at the 1% level. Most notably, a projection bias in satisfaction and hence also in loss aversion is absent when eliciting the return satisfaction independently from the investment decision in experiment III, which involves a simple risky asset with only six possible returns.

Even though convincing evidence of the projection bias is absent at the aggregate level, I use individual differences in satisfaction to test experience sampling as a possible cure. Experience sampling, a visual simulation of the possible return outcomes of a risky asset, has previously shown to considerably change investor behavior (Kaufmann and Weber (2013) and Bradbury et al. (2019)). I find no significant effect of experience sampling on the projection bias in the few ranges of investment returns in which I observe a significant difference in satisfaction. Experience sampling even seems to reduce anticipated satisfaction rather than increasing it to close the gap to the realized satisfaction.

Finally, I analyze the relationship between investment behavior and satisfaction with returns. A successful (failed) investment strategy increases (decreases) satisfaction with realized return outcomes only if the initial allocation is presented prominently next to the investment outcome when participants state their subjective ratings on satisfaction over realized returns. The regression estimate for the effect of the initial allocation on realized satisfaction is statistically significant in experiment II, but not in experiment III. But even if the information on the allocation is provided, and therefore the failure or success of a decision is made salient to the participant, investment success does not influence future investment behavior. For participants that realized a loss, differences in satisfaction, e.g. learning that losses loom larger in anticipation than after their realization, also does not affect future investment decisions. The importance of the salience of investment success suggests that even if a projection bias in satisfaction or loss aversion existed, investors might not adapt their investment behavior until they are (made) aware of their bias.

Overall, I conclude that previous findings on the projection bias in loss aversion (Merkle (2020)) and in satisfaction with outcomes (Kermer et al. (2006)) have to be interpreted with care. While previous work has estimated loss aversion coefficients on aggregate data separately for anticipated and realized returns across individuals, I study differences in satisfaction with the same outcome when evaluated ex-ante (anticipated) versus ex-post (realized) within the same individual. This granular view allows a more detailed investigation of the root of the projection bias and its implications for investment behavior. Kermer et al. (2006) elicit subjective ratings of satisfaction in a computer-simulated card-gambling experiment, and Merkle (2020) elicits investor satisfaction in a panel survey among clients of a large UK bank. Both papers find that people anticipate that they will feel emotional distress from suffering losses, but realized losses turn out to be less painful than expected. A possible explanation is that people tend to overestimate the hedonic impact of losses and underestimate their ability to cope with them. Consequently, the predicted hedonic impact attributed to losses is too large, and when people actually experience a loss, the realized impact on utility is smaller than the prediction. The "forecasting error" in anticipating satisfaction with an outcome in Kermer et al. (2006) is likely driven by differences in the level of active involvement between the treatment and control groups. Also, the "loss aversion illusion" found by Merkle (2020) might be biased because satisfaction is measured based on unrealized paper losses and remembered utility rather than realized losses and instant utility (see Imas (2016), Kahneman et al. (1997)). It is yet unclear how the satisfaction with return outcomes, and especially with losses, evolves over time after the return realization. Intuitively the hedonic impact should fade out. Furthermore, Merkle cannot rule out selection effects. Estimates for anticipated returns and perceived values for realized returns are provided endogenously together with the corresponding satisfaction by the survey participants. This is further aggravated by the fact that loss aversion is correlated with risk taking and thus determines investment success, which can significantly influence satisfaction with realized returns, as I show in this paper. I propose gathering more evidence before demoting loss aversion from preference to behavioral bias (see Merkle (2020)).

I contribute to the literature on loss aversion (Kahneman and Tversky (1979)) and its estimation (see among others Tversky and Kahneman (1992), Abdellaoui et al. (2007), Abdellaoui et al. (2008), Booij and van de Kuilen (2009), Booij et al. (2010)) and more specifically on loss aversion among private investors (e.g. Benartzi and Thaler (1995)). My research further adds to the literature on emotions in (financial) decision making (see e.g. Shiv, Loewenstein, and Bechara (2005), Shiv, Loewenstein, Bechara, et al. (2005), Kőszegi and Rabin (2008)) and projection biases (e.g. Loewenstein et al. (2003), Loewenstein (2005), Conlin et al. (2007), Acland and Levy (2015)) as affective forecasting errors (Wilson and Gilbert (2003)) in the prediction of future utility. Finally, I add to the discussion on a projection bias in loss aversion (see Imas et al. (2016) on the anticipation of loss aversion and Kermer et al. (2006) for experimental evidence) that has recently also been studied among private investors (Merkle (2020)).

The remainder of this paper is structured as follows. Section 1 reviews the related literature. Section 2 introduces the methodology and research questions. Section 3 presents all three experiments and their results. Finally, in section 4, I summarize our findings and then critically assess the results of previous research and conclude.

1. Literature review: Utility, satisfaction, and the projection bias

I begin this section with an in-depth review of the literature on the relationship between subjective satisfaction and utility, to emphasize the importance of measuring the hedonic impact of an investment outcome as well as observing the allocation decisions in an investment task. Furthermore, I summarize the literature on projection biases in the context of utility, emotions, and decision making. Utility is an abstract term that origins from early economic models to describe the benefit received from a certain level of consumption. In multi-period models with portfolio savings, utility additionally depends on consumption in the current period and the future consumption opportunities resulting from investing in risky assets while balancing their risk and return (see, for example, Merton (1969) and Samuelson (1969)). Modern theory argues that utility entails much more than what observed choices of consumption and saving can reveal about consumer preferences (Kahneman and Snell (1992), Kőszegi and Rabin (2008)). This gap is targeted by research on emotions, happiness, and satisfaction.[2] Traditional models assume a rational cost-benefit analysis that relies on objective probability weighting over possible outcomes. But this view has been rendered unrealistic. For example, anticipated emotions such as regret and disappointment have been argued to affect the decision making process (see, for example, Loomes and Sugden's (1982) model of regret).

Emotions experienced at the time of a decision can also affect decision making. First, "mood misattribution" happens when "mood tends to inform decisions even when the cause of the mood is unrelated to the decision being made" (Dowling and Lucey (2003)). For example Saunders (1993), Cao and Wei (2002), Hirshleifer and Teoh (2003), and Kamstra et al. (2003) have studied the impact of seasons and weather, as a proxy for mood, on stock prices. Goetzmann and Zhu (2005) and Schmittmann et al. (2015) studied the impact on trading behavior of individual investors directly.

2. Dowling and Lucey (2003) provide a good summary of the effect of emotions and feelings on investor decision-making.

Second, for the "risk-as-feelings" model, Loewenstein et al. (2001) argue that "every aspect of the decision-making process is influenced by the feelings of the decision-maker" (Dowling and Lucey (2003)). This means that while cognitive evaluations induce emotions, emotions inform cognitive evaluations, and finally, emotions affect behavior.

Emotions, mood, and feelings are thus essential to explaining human decision making. But measuring the feelings of investors during the investment process has so far received little attention in finance. Grosshans and Zeisberger (2018) elicit satisfaction on price paths of simulated stock investments and find that a stock catching up on previous losses (down-up price path) is much more attractive than a stock that recently lost part of its previous gains (up-down price path), even though they achieved the same return over a 12-month period. The authors furthermore elicit selling propensities for the different price paths in their experiments, but they do not directly examine the impact of the stated satisfaction on the investment behavior. Merkle et al. (2015) use a panel survey among clients of a large UK bank to assess the return levels for anticipated satisfaction and the satisfaction on realized past returns. Originally, Merkle et al. (2015) and Merkle (2020) use the term "happiness" rather than satisfaction, which I substitute to avoid confusion. Merkle et al. (2015) find that anticipated satisfaction is driven by return expectations, i.e. the higher the satisfaction, the higher is the required lowest return an investor would still be happy with. Anticipated satisfaction is furthermore influenced by the investment horizon, investment objectives and the portfolio risk taken. The authors thus conjecture that "aspiration" is a major component of the relationship between return and anticipated satisfaction. Realized satisfaction on the other hand is most strongly influenced by the actual past performance, but relative performance and active trading success also have a significant impact.[3]

The projection bias on utility

Even without external factors like weather, anticipated emotions, emotions experienced at the time of decision making, and emotions after the realization of an event can diverge significantly. This phenomenon is called the projection bias and emerged from research on the hedonic impact of choices. A series of papers introduced the idea that the "anticipated utility" or "decision utility" used in traditional models of rational choice does not necessarily coincide with the "instant utility" recorded in the moment of experiencing an outcome. Research started out by exploring and defining different concepts of utility, testing the validity of hedonic forecasting, and exploring the implications of subjective ratings of experienced utility in economic contexts.

Kahneman et al. (1997) provide further arguments against relying on decision utility alone. They counter the view that "subjective hedonic experience cannot be observed" or that "choices reveal

3. Merkle et al. (2015) ask very distinct questions for anticipated and realized satisfaction. As mentioned before, using happiness rather than satisfaction, anticipated satisfaction is asked about relating to the returns over the next 3 months: A) "What is the lowest return you would still be happy with?" B) "What is the highest return which you would still be unhappy with?" Realized returns are elicited on returns over the past 3 months: "How would you rate the returns of your portfolio? On a seven-point scale from "extremely bad" to "extremely good". When comparing survey responses for anticipated and realized satisfaction, only rarely do the authors find inconsistencies. Hence, they conclude that anticipated and realized returns are well aligned and they cannot confirm a projection bias in the satisfaction with returns (Loewenstein et al. (2003)).

all necessary information about the utility of outcomes" because agents are rational. The authors focus on "remembered utility" as the stated hedonic impact of past events and the repeated measurements of "instant utility", i.e. "immediately reported hedonic values of a moment of experience", on outcomes of long duration.[4]

If people have difficulty evaluating the pleasure or pain of recent past events, it is not surprising that forecasting the hedonic impact of outcomes is even more challenging. Wilson and Gilbert (2003) coined the term "affective forecasting". They found that people are skilled in predicting which situations they would dislike or like and predicting the type of emotions, like anger or joy, they will feel in different situations. But people are less gifted at forecasting the intensity and duration of emotions. The authors conjecture that people are not able to account for their sense-making processes that "ordinize" an event, they are thus prone to overestimate the emotional impact of an event. Loewenstein et al. (2003) use a different term, calling the discrepancy between the prediction of future utility and the experienced utility a "projection bias". The authors discuss evidence for a "general bias in the prediction of future tastes" and claim that people "systematically underestimate the magnitudes" of changes in their tastes. They claim that people are in general very good at adapting to major changes in their life circumstances, but have been found to underestimate this ability: The hedonic evaluation of events such as suffering serious medical conditions, moving to different climates, or changing occupation have a more dramatic impact for people giving prospective predictions than for people making retrospective evaluations (after successfully adapting to the changes in circumstances). Further examples of projection biases provided by the authors are the endowment effect, which induces people to value an object more when they actually own it compared to imagining that they own it.[5] And a projection bias in hunger, motivated by the folk wisdom "never shop on an empty stomach", appears as people overestimate their future hunger when they are currently in a state of hunger. After the seminal paper of Loewenstein et al. (2003), who also provide a formal consumption model with projection bias about future tastes, a series of research articles have documented a projection bias in different real-life situations such as medical decision making (Loewenstein (2005)), in catalog orders (Conlin et al. (2007)), in car and housing markets (Busse et al. (2012)), and in gym attendance (Acland and Levy (2015)).

Loss aversion and the projection bias

Only recently, the idea emerged that loss aversion in financial decision making could at least partially be a projection bias as well. Much of the existing literature focuses on the role of experienced outcomes and loss aversion (e.g. Strahilevitz et al. (2011), Thaler (1990), Imas (2016))

4. They report the result of a painful medical procedure as an example: If patients undergoing a colonoscopy received a prolongation of the examination, in other words the instrument was left in their colon for a little longer, they reported a less unpleasant evaluation of the treatment as a whole compared to the control group. The temporal integral of disutility should have exceeded that in the control examination because the examination of the treatment group was simply longer in duration but equally painful. However, the less painful additional endurance resulted in a less negative hedonic evaluation of the examination.

5. In experiments, people predicted a lower selling price for a coffee mug when they only imagined owning it, compared to actually possessing one and then determining a selling price (see Loewenstein et al. (2003)).

and explaining the important empirical phenomena, such as the equity premium or stock market participation puzzle (e.g. Benartzi and Thaler (1995)). More recent work has started to put emphasis on the importance of satisfaction for investors in general (Merkle et al. (2015)) and for investment decisions in particular (e.g. Grosshans and Zeisberger (2018)).

As an integral part of Kahneman and Tversky's (1979) prospect theory, loss aversion has contributed significantly to modern financial research, especially to asset pricing and personal finance. Due to its discouraging effect on demand for risky assets, loss aversion for example provides a plausible explanation for the Equity Premium Puzzle (Mehra and Prescott (1985)), i.e. the abnormally high risk premia observed in equity markets.[6]

For the same reason, loss aversion is likely to contribute to the low stock market participation rates observed in many countries (Mankiw and Zeldes (1991), Haliassos and Bertaut (1995)). Active traders are possibly affected by loss aversion when holding on to losing positions too long and selling winning positions too early, resulting in what is known as the disposition effect (Shefrin and Statman (1985)). Conversely, loss aversion may even induce increased risk taking. When investors try to recoup losses suffered in the past, they might be tempted to increase risk in exchange for larger expected returns in the future, a behavior known from casino players who raise the stakes in the hope of quickly recovering previous losses (Imas (2016)).

Given the previously discussed research that has shown how poor people predict their own preferences and emotions, a projection bias in the emotional impact suffered as a result of financial losses would not be surprising and would fit very well with the general failure to anticipate their own adaptability or "ordinization" mechanisms as emphasized by Loewenstein et al. (2003) and Wilson and Gilbert (2003). Imas et al. (2016) provide evidence that people are indeed able to anticipate loss aversion. In an experiment, workers were offered loss contracts as a performance incentive, in other words bonuses that could potentially be lost if their firms underperform. Participants were more likely to accept the offer when their loss aversion was stronger. They thus leveraged awareness about their loss aversion, proving that they know about their bias.

By contrast, Kermer et al. (2006) find a projection bias in loss aversion in a gambling experiment based on self-reported satisfaction which is smaller after losses for an actively involved treatment group compared to a control group who watched a computer play and only imagined having played and lost. And Merkle (2020) finds a projection bias in loss aversion among investors at a large UK bank. He uses subjective ratings on satisfaction for anticipated and realized returns.[7] He uses a seven-point scale to elicit utility as subjective ratings on satisfaction with investment

6. In combination with mental accounting, or narrow framing (Kahneman and Tversky (1984)), loss aversion leads to myopic loss aversion (Benartzi and Thaler (1995)): An investor bias that combines the fear of suffering losses in portfolio positions with framing individual positions into isolated short-term "mental accounts".

7. Anticipated satisfaction is elicited with the question: How would you rate the returns you expect from your portfolio held with us in the next three months? Seven-point scale from "extremely bad" to "extremely good." Realized returns are elicited with the question: How would you rate the returns of your portfolio (all investments held with us) over the past three months? Seven-point scale from "extremely bad" to "extremely good." Anticipated returns are best-guess predictions for portfolio returns over the next three months asked within the same surveys and realized returns are either perceived returns, also asked as a survey question, or real returns over the most recent three-month period (see Merkle (2020)).

returns. In the terminology of Kahneman and Snell (1990) participants report their "decision utility", i.e. the anticipated satisfaction with returns at the time of choosing the risky asset allocation. Additionally participants report their "experienced utility", i.e. the realized satisfaction after observing the actual investment outcome. Merkle (2020) identifies a significant projection bias in loss aversion with an estimated anticipated loss aversion coefficient of 2.2 against an experienced loss aversion coefficient of 1.2. Furthermore, he reports that investors with higher anticipated loss aversion hold less-risky portfolios, implying that anticipated satisfaction does have a real effect on investment behavior. Investigating potential causes for the projection bias in loss aversion, Merkle (2020) reports that loss aversion decreases after real losses occur in comparison to loss aversion estimated on anticipated satisfaction, and investors with higher financial literacy and more investment experience are less likely to suffer from a loss aversion illusion, i.e. a projection bias in loss aversion.

2. Methodology and research questions

I want explore the projection bias in satisfaction and loss aversion that has recently been described by Kermer et al. (2006) and Merkle (2020) and test it with experiments in a financial investment context.

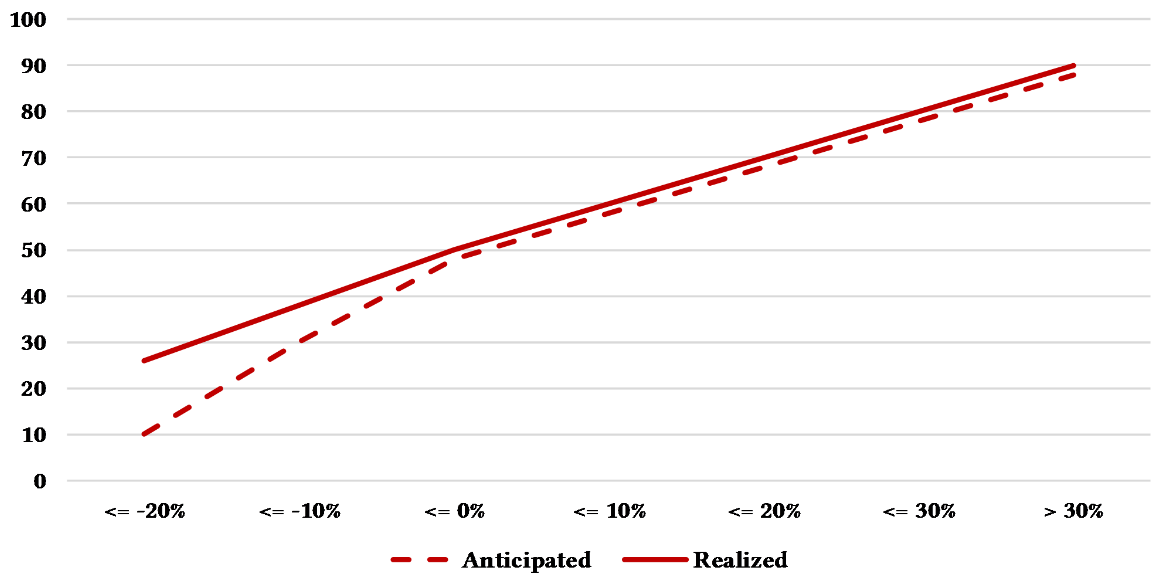


Figure 1.
Hypothesized shape of the satisfaction (utility) curves for anticipated and realized returns
The risky asset's returns are on the x-axis and the mean of elicitated satisfaction for a given return is on the y-axis. The curve for satisfaction on anticipated returns is dashed, for satisfaction on realized returns the curve is solid.

In Figure 1, I illustrate a hypothesized pattern of satisfaction for anticipated and realized returns. Loss aversion is calculated as the estimated change in utility over returns in the loss domain divided by the estimated change in utility in the gain domain. Utility curves are based on an elicitation of satisfaction with investment returns. The dashed line represents the utility estimated on investors' anticipated satisfaction with returns, the solid line represents the utility estimated on the realized satisfaction with returns, i.e. after actually experiencing the returns. Merkle (2020) finds a loss aversion coefficient of around 2 for anticipated returns that is significantly greater than loss aversion for realized returns, representing the projection bias in loss aversion. A loss aversion coefficient of 2 means that the slope of the satisfaction curve is twice as steep as the slope in the gain domain. The projection bias might, for example, be caused by coping mechanisms that help investors cope with financial losses, i.e. negative returns. Hence, if anticipated and realized satisfaction behave as hypothesized, in the loss domain, realized satisfaction should be greater and its slope less steep compared to anticipated satisfaction, as illustrated in Figure 1.

To clarify the role of loss aversion I define of the prospect theory value function as in Merkle (2020):

$$u(x) = \begin{cases} \eta x & \text{if } x \geq 0 \\ \eta \lambda x & \text{if } x < 0 \end{cases} \tag{1}$$

In Equation (1) $u(x)$ represents the utility that an investor receives for an outcome x, with a simple scaling factor η, and the loss aversion coefficient λ. The loss aversion coefficient represents the ratio of the slopes in a satisfaction-return diagram for the loss domain over the gain domain. Loss aversion is present for λ>1 given that losses have a larger hedonic impact than gains. I abstract from a reference point since I will only analyze percentage investment returns for the outcomes x. As in Merkle (2020), I simplify the original prospect theory utility function (Kahneman and Tversky (1979)) by estimating linear approximations for the relationship between outcomes and utility, abstracting from the diminishing marginal utility for larger gains and losses.

Research questions

In the following paragraphs I briefly outline my research questions and the corresponding hypotheses.

Can I verify the projection bias in loss aversion?

I want to explore the projection bias in satisfaction in a financial investment context that has recently been identified in card game experiments by Kermer et al. (2006). Thereby, I aim to verify findings by Merkle (2020), who finds a projection bias in the loss aversion estimated across investors by a measure of satisfaction elicitated during a recurring survey.
Accordingly, I test *hypothesis 1: There is a projection bias in loss aversion.*

In contrast to the aggregated estimation of the loss aversion coefficient as in Merkle (2020) and the comparison of happiness between two very distinct groups used in Kermer et al. (2006), I elicit anticipated and realized satisfaction for equal returns at the individual level. Coping mechanisms can explain why realized losses are less distressing than anticipated. Hence I expect that the

projection bias in loss aversion is driven by the projection bias in the satisfaction returns in the loss domain.
I test *hypothesis 2: There is a projection bias in anticipated versus realized satisfaction in the loss domain.*

Does experience sampling help against the projection bias?

Experience sampling, a visual simulation of possible return outcomes of a risky asset, has shown to considerably change investor behavior (Kaufmann and Weber (2013) and Bradbury et al. (2019)). After the information treatment, investors might choose riskier allocations, because they can better assess the riskiness of their investments. Given that Merkle (2020) reports that the projection bias decreases with more investment experience, I expect that anticipated satisfaction for losses increases for the experience sampling group.
I test *hypothesis 3: Experience sampling reduces the projection bias in satisfaction on investment returns.*

Does a projection bias influence investment decisions?

Merkle (2020) finds that experiencing portfolio losses mitigates the projection bias in loss aversion. If investors notice that realized losses are not as bad as expected they might increase risk taking.
I test *hypothesis 4: A larger projection bias leads to more risk taking in future investments.*

Do investment decisions influence the realized satisfaction?

Merkle (2020) notes that investment success is an important driver of investors' happiness, portfolio returns relative to peers and market performance significantly influence satisfaction. Zeelenberg et al. (1998) suggest that an active role in a decision making process increases satisfaction with outcomes. If participants take large bets, for example investing all of the endowment into the risky asset or none at all, satisfaction will certainly react more strongly to outcomes of the risky investment.
I test *hypothesis 5: The initial allocation influences the experienced return satisfaction.*

3. Experiments

General experimental flow

In a series of three experiments I analyze the projection bias in loss aversion and the underlying projection bias in anticipated versus realized satisfaction over investment returns and the impact on future investment behavior.

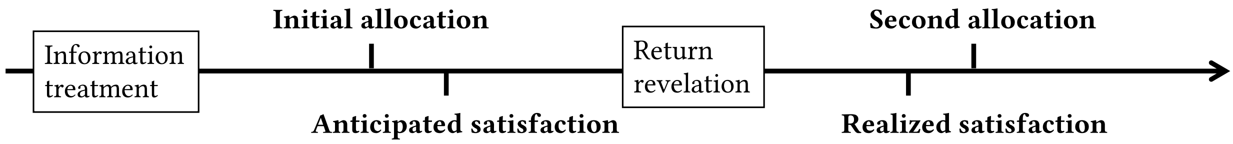


Figure 2.
Experimental flow
The flow shows the most important components of the experiments. Information provided to participants (in boxes) and values provided by participants (in bold text) marked by the ticks on the timeline. Note that only in experiment 3 the order between providing the initial allocation and the anticipated satisfaction was reversed.

In Figure 2, I illustrate the general experimental flow. At the core of all three experiments, participants in a first step decided how to allocate an endowment of $1,000 (€1,000 for experiment I and III) between a risky and a risk free asset, stating the *initial allocation.* Participants were instructed about the risky asset's return distribution using different *information treatments*; details are provided in the following sections. After the initial allocation, participants were asked to give subjective ratings, called *anticipated satisfaction* on a scale from 0 to 100 for six different potential returns from the risky asset's distribution. At the end of the experiment, a market simulation drew one random return out of the previously described return distribution, for each participant individually. Participants were informed about their investment return and then asked about their satisfaction with the realized return, the *realized satisfaction.* To explore how previous outcomes and anticipated and realized satisfaction affect risk taking, participants were then asked to make a further allocation of their endowment if they could invest for a second year, stating the *second allocation.*

Main differences between experiments

Before presenting the precise task, treatments, and results for each experiment, I summarize in detail their main objectives and differences. An overview table on the differences between the experiments can be found in Appendix A. The detailed running order of each experiment, the experimental instructions including the treatments are provided in Appendix C.

Satisfaction elicitation

All three experiments primarily aim at testing the projection bias in loss aversion and satisfaction. The comparison of anticipated and realized satisfaction is conducted based on a random return out of the risky asset's return distribution. The return to be realized for each participant was randomly drawn before eliciting the anticipated satisfaction. I asked for the anticipated satisfaction with six fixed returns. The realized return appeared first in the elicitation task to avoid anchoring. However, the random realized return will not necessarily exactly match one of these six anticipated returns. Therefore in experiment I, I interpolate anticipated satisfaction. In experiments II and III, I improve upon experiment I. In experiment II, the returns for which I asked anticipated satisfaction were held constant between subjects except one. Out of the six returns, I replaced the closest one with the later realized return. The risky asset in experiment III only had six possible returns, hence I always elicited the anticipated satisfaction on all possible returns including the randomly drawn return before its realization. This enabled me to test the root of a potential projection bias, i.e. differences in anticipated and realized satisfaction at the investor level without any deception and with having a return draw out of the described return distribution. In the terminology of Kahneman et al. (1997) I elicit realized satisfaction in the sense of "instant utility" after return realization in contrast to "remembered utility" as in Merkle (2020), who uses a quarterly survey that may not pick up the instant hedonic impact of portfolio performance, but rather the memory of the hedonic value. We keep in mind that the memory of hedonic events might be just as poor as their prediction, given that they are easily influenced and change over time (Kahneman et al. (1997)).

Experimental order and allocation decisions

In experiments II and III, I added an array of comprehension questions before the relevant decision tasks to increase the general understanding. Experiments I and III used a break of several days to separate the elicitation of anticipated and realized satisfaction, while experiment II was conducted in one session with literacy and demographic questions as an intermediate section between anticipated and realized satisfaction statements. In contrast to experiments I and II, in experiment III, the anticipated satisfaction statements are made before the first investment decision, and the realized satisfaction statement is made in the second part without displaying the initial allocation made in the first part. Furthermore, I incentivized the second investment decision to improve the evaluation of the impact of satisfaction with past returns on investment behavior.

Treatments and secondary objectives

Each of the three experiments had a special focus: In experiment I, I explore whether differences in anticipated and realized satisfaction are driven by the complexity of the return distribution and the way investors are informed about the risky asset. Participants were randomly assigned to a simple risky asset with two possible outcomes, -10% and +20%, with equal probability to assess the impact of complexity.

In experiment II, I used two different presentation formats for the information treatment that were also randomly assigned to participants.[8] For my *descriptive treatment*, I used a distributional graph to communicate riskiness via probability statements. I included reading examples, and a piece-wise introduction of the graph in order to facilitate understanding. For my *experience treatment*, subjects were informed with the help of an experience sampling procedure. More specifically, they were shown 30 draws of returns of the complex risky asset (16 draws for the simple risky asset in experiment I) based on the underlying distribution. I tested both presentation alternatives, as for example Kaufmann and Weber (2013) and Bradbury et al. (2019) have shown that experience sampling procedures lead to a better understanding of the risk-return framework relative to a classic descriptive presentation format.

In experiment III, I aim to measure satisfaction with a focus on returns while, in contrast to the previous experiments, minimizing the impact of investment decisions, in particular the initial allocation, on the anticipated satisfaction. To this end, in contrast to experiments I and II, participants stated the *anticipated satisfaction* before the *initial allocation*. Furthermore, participants were not reminded of the allocation they had made – not when the return on their investment was realized, not when they stated the *realized satisfaction*, and not when they decided on the second investment, the *second allocation*.

Risky assets

Experiment I included a simple risky asset with two possible outcomes and a complex risky asset with a continuous distribution. Experiment II used only a complex risky asset, identical to the complex risky asset in experiment I, whereas experiment III used a risky asset with six possible return outcomes of equal probability illustrated as a die.

Participants

For experiments I and III, I recruited economics students from Radboud University (The Netherlands). Participants for experiment II were regular people acquired via Amazon's Mechanical Turk.

3.1. Experiment I – Risky asset complexity

Task

Participants were first asked to allocate an endowment of €1,000 between a risky and a risk free asset by selecting the amount to be invested in the risky asset (the *initial allocation*) on a slider with full units. The *complex risky asset* had an annual expected return of 8% with a volatility of 17%. Returns were sampled from the monthly MSCI World US$ time series from 1976 to 2016. Alternatively a *simple risky asset* with two equally possible outcomes, -10% and +20%, was provided. While not identical, the expected return of the simple asset was of similar magnitude to the expected return of the complex risky asset. The risk free rate was set to a zero return. Before the allocation, participants received the information treatment, illustrating the return distribution

8. In experiment I, I used the same set of information treatments, but the number of observations was not sufficient for a conclusive comparison.

of the respective risky asset and were asked about their return and risk perception. Details on the information treatment are described in the following paragraph on the treatments in the experiment. Next, participants stated their *anticipated satisfaction* on a slider from 0 to 100 for six different potential returns earned on their investment. These returns corresponded to the 10th, 25th, 40th, 55th, 70th and 85th percentiles of the risky asset's distribution. After stating their anticipated satisfaction, participants were asked to state their risk preferences, answer a financial literacy questionnaire, and provide information on their demographics. See Appendix C for a transcript of the experiment instructions. The experiment was split into two parts to avoid an anchoring effect of anticipated satisfaction on realized satisfaction. In the first part, the participants stated their anticipated satisfaction and made an investment decision based on the return information treatment. Two weeks after the initial survey experiment, in the second part, a "financial market simulation" drew one random return out of the previously described return distribution, for each participant individually. Participants stated their *realized satisfaction* with the return observing their invested amount, i.e. the *initial allocation*. Finally, participants provided the *second allocation*, the amount between €0 and €1,000 they would allocate to the risky asset if they could invest a second time after observing the *realized return*. See Figure 2 and Appendix C for details of the experimental flow.

Treatments

Experiment 1 served the purpose of testing the impact of return complexity and different information treatments on the satisfaction with returns in anticipation and after realization as well as the resulting projection bias. Participants were randomly assigned to the *simple risky asset* or the *complex risky asset*. I used two different information treatments. In the *descriptive group*, participants were shown a distributional graph to communicate riskiness via probability statements. In the *experience group*, participants were informed with the help of an experience sampling procedure. More specifically, they were shown 30 draws of returns for the complex risky asset (16 draws for the simple risky asset) based on the underlying distribution. See Appendix C for examples of the information treatments. In the experiment participants were first shown either a descriptive graph or experience sampling of random returns from the distribution of the risky asset. Next, they stated their allocation and the anticipated satisfaction. The information treatment was repeated at the beginning of the second part of the experiment before the market simulation revealed the realized return.

Participants

Experiment 1 was conducted with undergraduate students in economics from Radboud University, Netherlands in October 2016 with 318 participants, see Table A.1: 36.2% of students were female, 9.1% stated that they owned stocks or mutual funds. The average duration was 20 minutes for part one and 5 minutes for part two. The average duration was 20 minutes for part one and 5 minutes for part two. The average amount intially allocated to the risky asset was €580.50, the investment outcome (=risky allocation * (1+risky asset return)) was €625.80, and the final outcome was €1045.29 on average. I picked every 10th participant at random, independently of their decisions, to receive the final outcome of the experiment divided by 100 as their payoff. The actual (expected) payment was €10.45 (€1.05) on average for 25 minutes which corresponds to €25.08 (€2.51) per hour. The actual average payment is the average amount paid out to participants actually selected for the payout. The expected average payment is: (the average risky allocation*(1+the expected risky return) + the riskless allocation)/100 (payout factor)/10 (probability to be drawn for payout).

Results

The simple asset had only two possible outcomes, a loss of -10% and a gain of +20% with equal probability. Losses received an anticipated (realized) satisfaction of 26.7 (23.5), whereas gains reached 79.2 (78.1). I find no significant difference between the satisfaction with anticipated returns and realized return. In the loss domain, i.e. for a return of -10%, the anticipated satisfaction for the description group is, on average, even 6.3 points higher than the realized satisfaction, but the difference is not statistically significant. See Figure 4 and Table B.2. Hence, I focus on the complex risky asset to test my hypotheses.

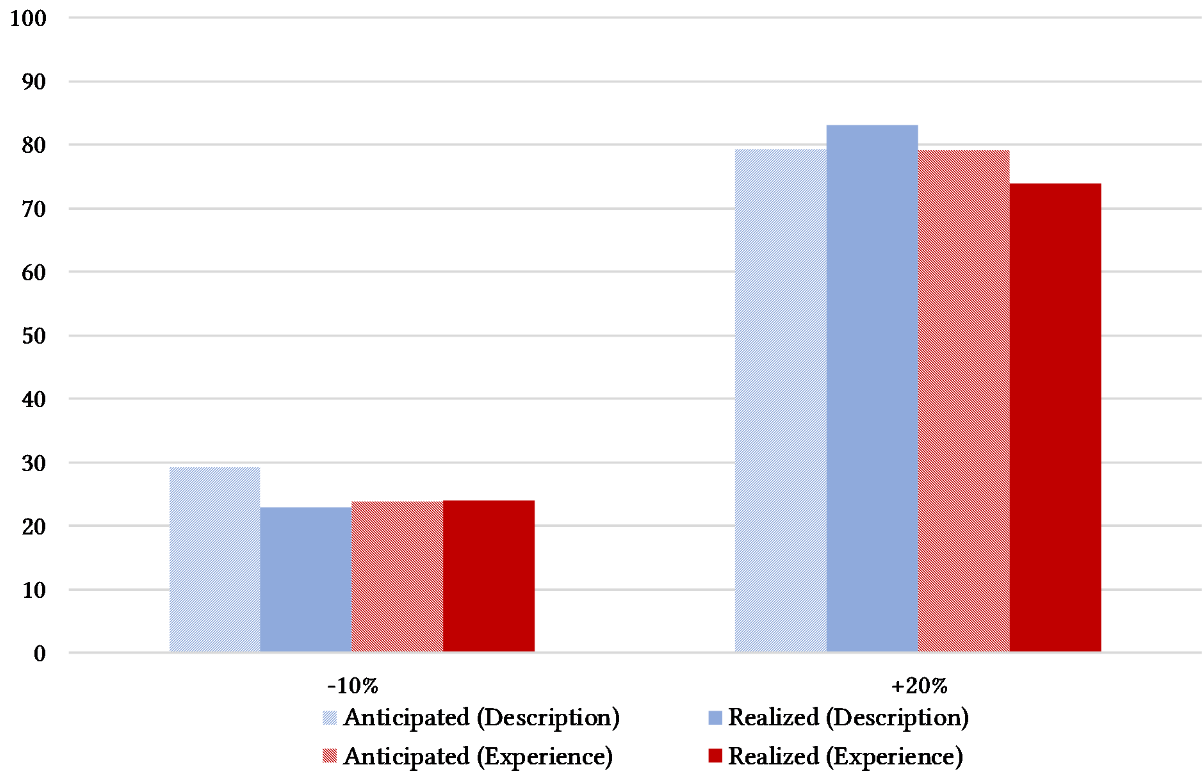


Figure 3.
Satisfaction with returns for the simple risky asset (experiment 1)
Bar chart on satisfaction averages by information treatment. The experience sampling (description) treatment is colored in red (blue). Realized (anticipated) satisfaction is indicated by solid (textured) bars.

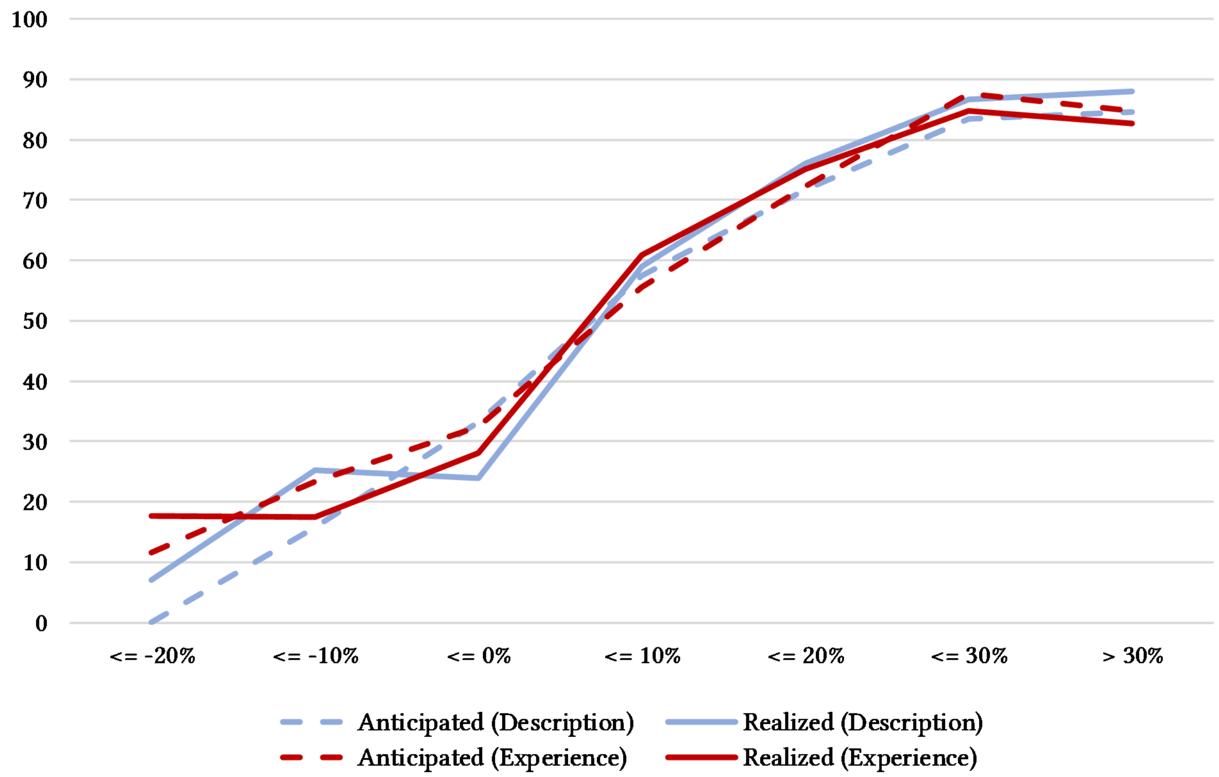


Figure 4.
Satisfaction with returns on the simple risky asset (experiment 1)
Line chart on satisfaction averages by information treatment. The experience sampling (description) treatment is colored in red (blue). Realized (anticipated) satisfaction is indicated by solid (dashed) lines.

Hypothesis 1: There is a projection bias in loss aversion.
I estimate loss aversion coefficients following Merkle (2020) using only participants of the complex risky asset group, since the two outcomes of the simple asset do not allow an estimation of slopes for the impact of investment return on satisfaction in the loss and gain domain separately. The loss aversion coefficient is estimated at 1.15 for anticipated returns and at 1.18 for realized returns. The two coefficients are not statistically different from each other. Notably, they are far lower than the loss aversion coefficients found in the previous literature (see e.g. Merkle (2020)) and do not show evidence for loss aversion since both coefficients are not statistically different from 1. See B.1, column 1 for details on the regression results.

Hypothesis 2: There is a projection bias in anticipated versus realized satisfaction in the loss domain.

For the complex asset, I calculate the average satisfaction on intervals of 10% of its continuous return distribution (except of the first and last, which cover the tails of the return distribution). With 216 participants in the group with the complex risky asset, and an additional split between the description and experience information treatment, I have at most 31 participants in each of the 7 intervals, see Table B.2 for the number of observations, means for each interval and treatment group. In the gain domain, anticipated and realized satisfaction seem very close. In the interval of returns larger than 0% and smaller or equal to 20%, the average realized satisfaction is larger than the average anticipated satisfaction. The difference is significant for the interval of returns larger than 10% and smaller or equal to 20% at the 5% level.

I define the projection bias in satisfaction as the difference between realized satisfaction and anticipated satisfaction. I expected a positive projection bias in satisfaction in the loss domain (hypothesis 2). My results are mixed. For the description sample, I find a positive and significant projection bias in the interval for returns smaller or equal to -20%, but the result relies on a single observation. The projection bias for the experience treatment is positive and of a similar magnitude but not statistically significant. In the interval for returns larger than -20% and smaller or equal to -10%, the description treatment shows a large positive projection bias of 9.4 and the experience treatment a negative projection bias of -5.9 which is most likely driven by the high anticipated satisfaction in the experience treatment. For returns between -10% and 0%, the projection bias is negative. This might be due to the fact that anticipated satisfaction for realized returns is estimated by interpolation. With anticipated satisfaction reported for the two returns -10% and 0%, it is very likely that the anticipated satisfaction for returns equal to 0 biases upward the estimates for anticipated satisfaction on negative returns between -10% and 0%. At 0% there might likely be a jump in satisfaction that is not captured by my discrete elicitation and would require capturing satisfaction at a small negative return to appropriately interpolate on the interval -10% and 0%. In experiments II and III, I managed to elicit anticipated satisfaction with returns that match the realized returns and reduced the dimension of treatments by removing the simple risky asset. I abstain from testing hypothesis 2 using experiment I in a regression setup, since the estimation of anticipated returns using interpolated values is likely to bias the results.

Hypothesis 3: Experience sampling reduces the projection bias in satisfaction on investment returns.

The anticipated satisfaction with returns in the loss domain is larger for the experience treatment group compared to the description group. This is in line with hypothesis 3. I expected experience sampling to increase the anticipated satisfaction in the loss domain. Testing this hypothesis in an OLS regression, I find a c.p. increase of 8.5 on the anticipated satisfaction for experience sampling in the group for the complex asset compared to the simple asset – description treatment, see Table B.3 column 4. Compared to the complex asset – description treatment, the effect of 3.06 (adding the estimated coefficient of -5.04 on the dummy for the complex asset) is still positive but all estimates are insignificant. In the gain domain, in the case of the simple risky asset, experience sampling decreases the realized satisfaction by 8.58 points, which is significant with a p-value of 4.87%. But the effect is reduced to 1.11 in the case of the complex asset. Overall, I see a tendency for experience sampling to correct a potential projection bias in the loss domain, but the corresponding estimates are not statistically significant enough to confirm the hypothesis. In experiment II, I revisit this hypothesis using a larger sample.

3.2. Experiment II – Information treatment: The role of experience sampling

Task

The setup in experiment II differs from experiment I in a few, but important, aspects. Experiment II was designed to calculate the projection bias without interpolation or extrapolation of the anticipated returns. To have a clean comparison between the anticipated and the realized satisfaction for each participant for exactly one return, the returns for which I asked anticipated satisfaction were held constant between subjects except one. One return was replaced with the return to be drawn as the realized outcome for each participant. Out of the six returns I replaced the one which was closest to the later realized return. The realized return appeared first in the elicitation task to avoid anchoring. The display order of the remaining five returns was randomized. As in experiment I, participants were asked to allocate an endowment of US$1,000 between a risky and a risk free asset and stated their *anticipated satisfaction* after the information treatment. In contrast to experiment I, the experiment took place in a single session. The break between the first and second part of the experiment was reproduced by asking participants to answer the questionnaires on demographics and numerical literacy between the initial allocation decision and the draw of the first return. Participants were informed that the following "financial market simulation" represented one year as in the previous experiment. As before, participants stated their *realized satisfaction* and the decision on the *second allocation* after the return of the risky asset was drawn. To ensure that participants clearly understood the experimental task, I added examples and comprehension questions at the beginning of the experiment. Details on the experimental flow and differences between experiments can be found in Appendix C.

Treatments

In experiment II, I removed the simple risky asset. The two information treatments remained the same, except that I included reading examples, i.e., a piece-wise introduction of the distribution graph, in order to facilitate understanding.

Participants

Experiment II was conducted on the Amazon Mechanical Turk with 427 participants, see Table A.1: It took participants an average of nine minutes to complete. 36.2% of participants were female, participants were, on average, 35.7 years old and 72.2% held a college degree. I added comprehension questions and improved the descriptive treatment to maximize savviness. Participants received a fix participation fee of $0.51. Additionally, as a variable incentive, I again picked every 10th participant at random, to receive the final outcome of the experiment divided by 100 as a bonus payment. The average allocation was $582.64, the investment outcome was $606.14, and the final outcome was $1023.61 on average. I picked every 10th participant at random, independently of their decisions, to receive the final outcome of the experiment divided by 100 as the payoff. The actual (expected) payment was $10.24 + $0.51 = $10.75 ($1.53) on average for 16 minutes which corresponds to $40.31 ($5,74) per hour.

Results

Anticipated (realized) satisfaction in the loss domain was, on average, 18.7 (22.6) for the descriptive treatment group and 16.6 (20.0) for the experience sampling group. In the gain domain, satisfaction reached, on average, 78.1 (77.7) for the descriptive treatment group and 72.6 (73.3) for the experience sampling group. I observe in Figure 5 that realized satisfaction closely follows the anticipated satisfaction, but that, in general, satisfaction seems to be slightly greater for realized returns than for anticipated returns, with some exceptions for higher returns. Interestingly, the line graph demonstrates why it is important to analyze satisfaction at the investor level to assess a projection bias in loss aversion. If investors are simply more pessimistic about future outcomes independently of gains and losses, the satisfaction curves would only shift vertically. This narrative seems to fit the graph better, since I do not see an upward shift in the slope for realized returns in the loss domain (compared to Figure 1).

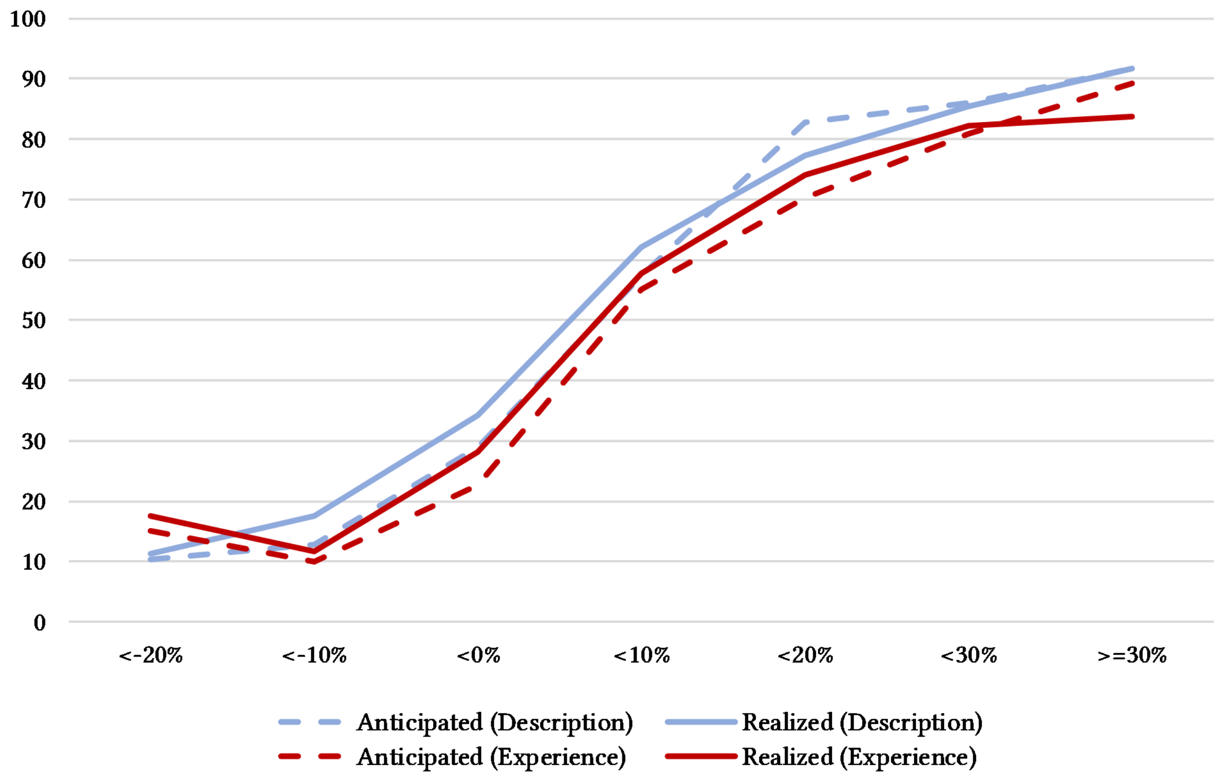


Figure 5.
Satisfaction with returns (experiment II)
Line chart on satisfaction averages by information treatment. The experience sampling (description) treatment is colored in red (blue). Realized (anticipated) satisfaction is indicated by solid (dashed) lines.

Table B.4 reports the average satisfaction over return ranges, for anticipated and realized returns and broken down by information treatment. Significance in the projection bias, i.e. the difference between realized and anticipated satisfaction, is tested using two-sided t-tests. I find a significant projection bias of 3.6 for the loss domain at the 1% level (see Table B.4, column 8, row named

Realized-Anticipated, *Overall*). In the ranges between -10% to 0% and 0% to 10% the differences are 5.4 and 3.7 (see Table B.4, columns 3 and 4, row named *Realized-Anticipated*, *Overall*) and significant at the 5% level.

Hypothesis 1: There is a projection bias in loss aversion.

The OLS regression estimates for the loss aversion coefficients of experiment II are reported in Table B.1, column 2. λ_{Antic} is the loss aversion coefficient for anticipated satisfaction on returns and λ_{Real} is the coefficient for realized returns. Surprisingly, loss aversion is 0.904 for anticipated satisfaction, thereby smaller than the estimate of 1.121 for realized satisfaction. The fact that it is actually smaller than 1 means that there is no evidence of loss aversion. Furthermore, realized loss aversion is estimated to be significantly different, i.e. greater than anticipated loss aversion with a p-value of 6.2%. I consequently have to reject hypothesis 1, a projection bias in loss aversion in the expected direction.

Hypothesis 2: There is a projection bias in anticipated versus realized satisfaction in the loss domain.

After providing evidence of a projection bias in satisfaction for the loss domain in simple differences, I discuss the results of multivariate regression tests. Table B.6 reports the OLS estimates for regressions of the differences in satisfaction (realized-anticipated) primarily on a dummy for the loss domain. Column 1 reports an overall difference in satisfaction of 1.762. This difference is driven by the loss domain. The estimate for the dummy on the loss domain is 3.312 and significant at the 5% level, see column 2. In the model reported in column 3, I additionally control for the risky allocation chosen by the participants, its interaction with the loss domain dummy, and the set of variables capturing demographics, sophistication, beliefs, and preferences. The coefficient on the loss domain dummy increases considerably to 14.3 and becomes significant at the 1% level.

Hypothesis 3: Experience sampling reduces the projection bias in satisfaction on investment returns.

In Table B.5, I show the results of an OLS regression for satisfaction on a dummy for the experience sampling information treatment, a loss domain dummy, and their interaction. Control variables include the usual set of participant characteristics and the actual level of return as well as its interaction with the loss domain dummy. To test whether experience sampling ameliorates the projection bias in satisfaction, I first test the effect of experience sampling on anticipated and realized satisfaction separately. Column 1 shows the regression results using anticipated satisfaction as the dependent variable. I find that experience sampling decreases anticipated satisfaction by 5.9 points in the gain domain, and the effect is significant at the 5% level. In the loss domain, this effect is reduced by 3.6 points but the difference in the impact of experience sampling between the loss and gain domain is insignificant. For realized satisfaction (column 2), the effects are much smaller: -3.6 for gains and $-3.6 + 0.6 = 3$ for losses, but the coefficients are insignificant. One potential explanation is that enduring the rather lengthy experience sampling process did affect the overall happiness of participants which was mirrored in statements on anticipated satisfaction that followed directly after. The effect might have worn off during the experiment, explaining why it is not as noticeable for realized returns. Overall, experience sampling decreases satisfaction in the loss domain slightly more for realized satisfaction than for anticipated satisfaction. This counts towards a rejection of hypothesis 3. Following Merkle (2020), who finds that experienced losses reduce the projection bias in loss aversion, I expected that experience sampling might increase satisfaction in the loss domain by familiarizing participants with the risky asset's possible outcomes. Hypothesis 3 is tested directly on the difference in satisfaction as the dependent variable (see columns 3 and 4). Experience sampling does reduce the projection bias by about 1 point in the loss domain, but the effect is small and all related coefficients are insignificant. Including or excluding the level of the risky asset's return, including its interaction with the loss domain dummy, does not affect the estimates considerably (compare columns 3 and 4). Hence I cannot find a significant effect of experience sampling on the projection bias and must reject my hypothesis.

Hypothesis 4: A larger projection bias leads to more risk taking in future investments.

The OLS estimates for the regression of the change in allocation (realized - anticipated) on the projection bias (realized - anticipated satisfaction) are reported in Table B.7. I start in column 1, regressing the change in allocation on only a dummy for the loss domain. Suffering losses seems to discourage future risk taking by about $207, significant at the 1% level. In column 2, I add the participant-specific projection bias in satisfaction and its interaction with a loss domain dummy. The results are surprising. A larger projection bias means either or both higher realized satisfaction and smaller anticipated satisfaction. An increase in the projection bias of 1 point leads to a decrease of 2.4 in the change of the allocation, but the effect is only significant at the 1% level. At the bottom of the table I report "Effect for Losses" which is a t-test for significance for the linear combination of the first two coefficients: Adding the coefficient on the satisfaction difference and its interaction with the loss domain estimates the effect of a larger projection bias on the difference in allocation for the loss domain. This effect of 0.124 is close to zero and insignificant. I must note that even the effect size for the gain domain is very small. Given that the mean difference in satisfaction never exceeded 6 points over all return ranges (Table B.4), a 6 point estimate for the projection bias would only induce an increase in the second allocation compared to the initial allocation of $15. Adding the set of participant control variables does not affect coefficients considerably (column 3), but adding the initial allocation and its interaction with the loss domain (column 4) further decreases the estimates on the projection bias. For the loss domain, the sign is reversed but the effect is too small and insignificant to be meaningful. An interesting finding is the significant coefficient estimate for the initial allocation. Given that the initial allocation already enters the dependent variable directly, I use the second allocation as the dependent variable in column 6. I find that, on average, future risk taking increases by $0.66 for every $1 invested in the initial allocation and the effect is highly significant. The interaction with the loss domain is very close to zero and insignificant. This means that participants repeat their risk allocation strategy to a large extent. I might expect that the success of participants' investment strategy would impact future risk taking, for example that if losses are realized after high risk taking (such as investing the maximum of $1000), this would have discouraged similar risk taking in the future. But my estimates show no evidence of such a reaction.

Hypothesis 5: The initial initial allocation influences the experienced return satisfaction.

In columns 4 and 5 of B.6, I repeat the regression of satisfaction on the initial allocation and control variables using the anticipated and realized satisfaction directly as dependent variables instead of their difference. In column 5, I test whether the initial allocation influences realized satisfaction. The estimates suggest that the satisfaction with returns increases by 0.014 after realizing a positive return for each additional dollar that the participant invested in the risky asset. Given that the endowment was $1000, investing $100 more would lead to a 14 point higher evaluation in satisfaction. This increase is large when comparing the increase in average realized satisfaction of only 8.2 when moving from the return range <= 20% to <= 30% in Table B.4. Likewise, every additional dollar invested into the risky asset decreases the realized satisfaction by 0.023, resulting in a loss in satisfaction of 23 for an increase of $100. Column 4 shows that the allocation decision is not correlated with the anticipated satisfaction.
I confirm hypothesis 5, given the strong evidence and large economic effects.

3.3. Experiment III – Return satisfaction: Factoring out investment decisions

Task

In experiment III, I aim to measure satisfaction with a focus on returns, while minimizing the impact of investment decisions. Further, I introduce an incentivized second investment round that should improve the evaluation of the impact of satisfaction with past returns on investment behavior.

First, I introduce a single risky asset with six possible return outcomes and equal probabilities. Therefore, as in experiment II, the return drawn in the market simulation will be among the six returns for which participants have to state their anticipated satisfaction. Second, I measured *anticipated satisfaction* before the *initial allocation* to ensure independence of the anticipated satisfaction with respect to the investment amount. It is obvious that deciding on a high investment amount is likely to decrease the anticipated satisfaction on negative returns. As in experiment I, I divided the experiment into two parts. The second part began with a review of the information on the return distribution and the market simulation which included the return realization and the following statement of the *realized satisfaction*. Part 2 was accessible starting two days after part 1. In contrast to experiments I and II, the amount invested in round 1, the *initial allocation*, and the nominal outcome of the investment were NOT displayed and prominently conveyed to participants. Instead, participants were only confronted with the percentage return outcome. By concealing the investment amount after the market simulation, I hope to minimize the impact of the *initial allocation* on the *realized satisfaction*. It is, however, unclear what direction the effect of the investor being made aware of a (higher) allocation would take in terms of the realized satisfaction, even when considering only the loss region. A higher allocation might intuitively lead to a smaller realized satisfaction after a loss has been realized, because the decision to invest riskily is regretted. But Zeelenberg et al. (1998) suggest that individuals are more happy with outcomes if they actively made a decision compared to being subject to a random process that they were not able to influence or take part in. Providing the investment amount in experiments I and II might not only make participants more certain of their exact allocation decision, but also increase their realization that they played an active role in the outcome. I am hence not able to predict the impact of concealing the initial allocation during the assessment of the realized satisfaction, but prefer a measure of realized return that is less influenced by the previous allocation decision.

Finally, after the return was drawn and presented, participants were asked to give their *realized satisfaction* and to make a second allocation decision, the *second allocation.* Furthermore, participants were informed that the second allocation counted as a new and independent round of the investment game, and that their payoff would be determined by randomly choosing one of the two investment rounds and the corresponding market returns. This ensures that the *second allocation* was incentivized in experiment III to make valid statements about changes in investment behavior between the first and second allocation. See Appendix C for the details on the flow of experiment III, the changes in the information treatment, and the display of the market simulation outcome.

Treatments

The risky asset has six possible outcomes, the returns -20%, -10%, 0%, 10%, 20%, 30% have equal probability and are framed as the outcomes of throwing a die. See Appendix C for the information treatment in experiment III. Participants randomly received one of the six outcomes. Experiment III abstracts from any additional stimuli to maximize the power of estimating and testing the projection bias and its impact on investment decisions.

Participants

Experiment III was conducted with students from Radboud University, Netherlands in October 2017 with 402 participants, see Table A.1. Just as in experiment I, it was split into two parts to avoid an anchoring effect of anticipated satisfaction on realized satisfaction. In the second part, which was accessible two days later, participants observed the return realization, stated their satisfaction and made the allocation decision for a second incentivized investment period. 40.8% of students were female, 14.7% owned stocks or mutual funds. The average duration for part one was 23, and 26 for part two. The average allocation was €619.30, the investment outcome was €654.78, and the final outcome was €1035.45 on average. I again picked every 10th participant at random, independently of their decisions, to receive the final outcome of the experiment divided by 100 as payoff. The actual (expected) payment was €10.35 (€1.04) on average for 49 minutes which corresponds to €12.67 (€1.27) per hour.

Results

Figure 6 shows the satisfaction curves representing averages for anticipated (realized) returns with a dashed (solid) line. The corresponding values are reported in Table B.8. The two curves are very close to each other except for the highest possible return of 30%, for which the realized satisfaction exceeds the anticipated satisfaction by 4 points, significant at the 1% level for a two-sided t-test. Notably, the satisfaction increases around 12 points moving to the next higher return in the loss domain (including 0% return). Satisfaction increases by almost 30 points from 0 to 10%, by around 20 points from 10% to 20% and by less than 10 points from 20% to 30%. This means that losses have a larger impact than gains on satisfaction only if I compare a 10 percentage point increase in losses to an increase of equal magnitude in the gain domain for returns larger than 20%. Given this initial evidence, a projection bias in loss aversion and in the loss domain of satisfaction seems very unlikely. Nevertheless, I will provide the multivariate regression tests to repeat the test on my hypothesis for experiment III.

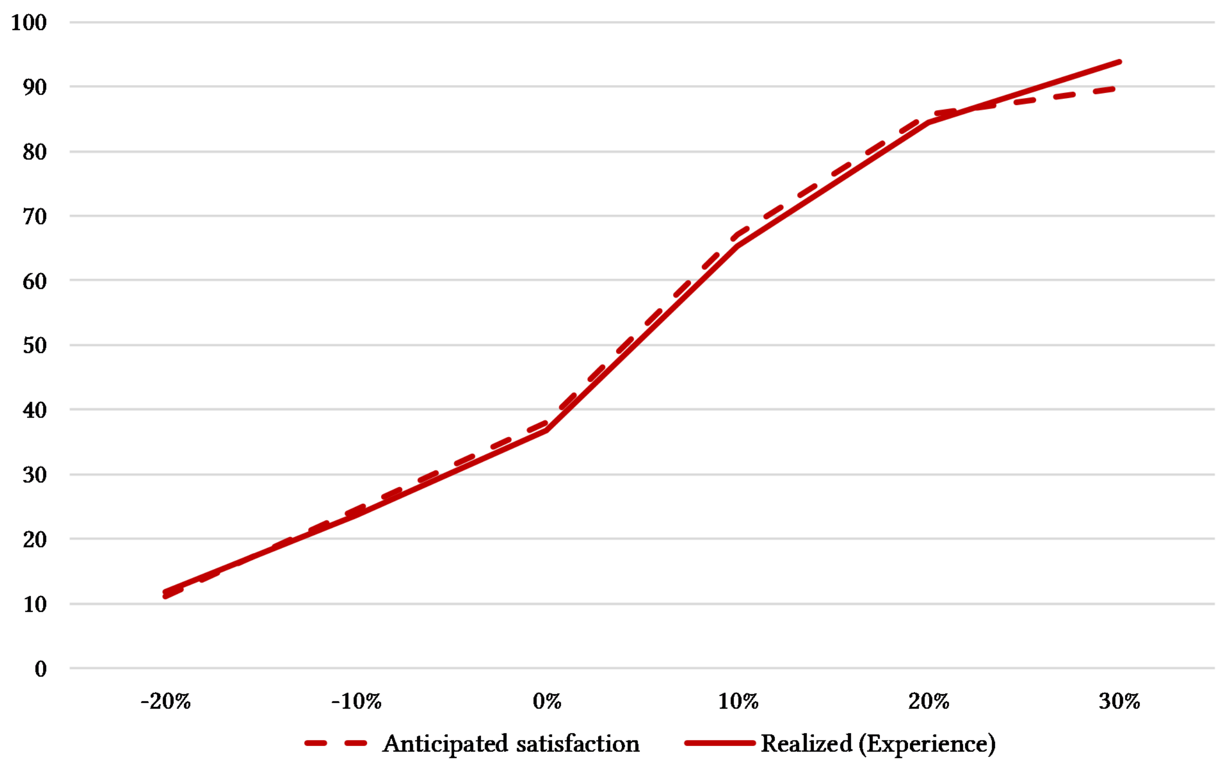


Figure 6.
Satisfaction with returns (experiment III)
Line chart on satisfaction averages. Realized (anticipated) satisfaction is indicated by solid (dashed) line.

Hypothesis 1: There is a projection bias in loss aversion.

The estimated loss aversion coefficients for experiment III are reported in B.1, column 3. In rows 1 and 2, we see that the slopes for realized returns are almost identical in the loss and gain domain. The coefficients in rows 2 and 3 suggest that the slope for anticipated returns, as compared to realized returns, is slightly steeper in the loss domain and slightly less steep in the gain domain, but only the difference in the gain domain is significant at the 1% level. In Figure 6, I noted the differences in satisfaction for returns of 30%, which likely drive this result. The finding is well reflected in the estimated loss aversion coefficients. λ_{Antic}, the loss aversion coefficient for anticipated satisfaction on returns, is estimated to be equal to 1.14, and the Wald test rejects the null of a loss aversion coefficient equal to 1 at a p-value of 4.2%. The estimated loss aversion coefficient for realized returns λ_{Real} is equal to 1.005 and not statistically different from 1. Hence there is evidence of small but statistically significant loss aversion in anticipated satisfaction, but none for realized satisfaction. I must reject the null-hypothesis of equal loss aversion (last row in Table B.1) at a p-value of 3.1%. Hence I cannot reject my hypothesis that there is a projection bias in loss aversion based on the data in experiment III. But loss aversion is small and driven by the surprise in satisfaction for the highest possible return. Furthermore, I note that using six discrete outcomes biases the estimate of the slopes for the gain and loss domain. If I categorize the return of 0% to the gain domain or generate two data points for zero returns (one for the loss and one for the gain domain), the sharp increase of 30 points between a return of 0% and 30% will increase the average slope in the gain domain so that loss aversion does not persist.

Hypothesis 2: There is a projection bias in anticipated versus realized satisfaction in the loss domain.

In Table B.9, column 1, I report the regression of the difference in realized - anticipated satisfaction on a constant only and a loss domain dummy. As expected from my previous results, I must reject the hypothesis of a projection bias in the loss domain. Adding participant controls has almost no effect on the coefficient on the loss domain dummy (column 2), and adding the initial allocation and its interaction with the loss domain dummy to control for success in the investment strategy (column 3) increases the absolute value of the estimate, but the difference is insignificant. The estimate is still small at 1.9 and implies a higher anticipated than realized satisfaction, counter to my intuitive argumentation of unanticipated coping mechanisms for painful losses.

Hypothesis 4: A larger projection bias leads to more risk taking in future investments.

Even though I cannot find realized satisfaction to be larger than anticipated satisfaction for negative returns, I test whether individual differences in satisfaction affect risk taking. Table B.10 reports the OLS estimates for the regression of the change in allocation (second - initial) on the projection bias (realized - anticipated satisfaction) for experiment 3. In column 1, I only regress on a constant and a dummy for losses and find that participants only decrease investments into the risky asset by €22 on average after suffering a loss in the first return draw, and the effect is not statistically significant. Testing hypothesis 4, I include the projection bias in satisfaction and its interaction with the loss domain (see column 2). For the gain domain, each point in the difference, i.e. surprise, in satisfaction increases the change in allocation by €2.2. But the effect is very small and only significant at the 10% level. For the loss domain, the effect is reduced to 0.335 and insignificant (reported at the bottom of the table). Including the set of participant controls and the initial allocation with its interaction with the loss domain dummy does not change the estimates considerably (see columns 3 and 4). As in experiment 2, I note that the initial allocation is a strong predictor of the second allocation, but the coefficient is only 0.44 compared to 0.66 in experiment 2. Again, I cannot find an effect for a successful or failed allocation strategy, given that the interaction of the initial allocation with the dummy on the loss domain is close to zero and insignificant.

Hypothesis 5: The initial allocation influences the experienced return satisfaction.

In column 5 of Table B.9, I examine whether differences in the initial allocation lead to higher satisfaction after receiving a return in the gain or loss domain. In contrast to experiment 1 and 2, I did not show participants the allocation after their decision on the risky investment. Apparently, this completely eradicates the impact of the investment decisions on the realized satisfaction; the coefficients on the initial allocation and its interaction with the loss domain are close to zero and statistically insignificant. In column 4, I confirm the same for anticipated returns, which is not surprising given that anticipated satisfaction was elicited before the initial allocation decision.

4. Summary and conclusion

In this section, I first summarize my numerous findings before concluding with a discussion in the context of previous findings of related research.

Loss aversion in anticipation and for realized returns

Estimated loss aversion for both anticipated and realized satisfaction is close to 1 for all experiments. I find a projection bias ($\lambda_{Antic} > \lambda_{Real}$) in loss aversion, with λ_{Antic} statistically different from 1 and λ_{Real}, only in experiment III. The estimate for $\lambda_{Antic} = 1.138$ is small compared to the estimate reported by Merkle (2020) of around 2, but all estimates on loss aversion are within the interquartile range of estimates documented by Zeisberger et al. (2012). I find that the loss aversion in experiment III is not driven by a surprise in satisfaction for outcomes in the loss domain. Participants stated, on average, a 4 point higher realized satisfaction than anticipated satisfaction after receiving a return of 30%. More importantly, the finding is not robust to adding the observations for satisfaction on the 0% return to the gain domain since satisfaction increases, on average, by almost 30% when comparing 0% and 10% returns (see table B.8).

Why are Merkle (2020)'s estimates on anticipated loss aversion so much larger than the realized loss aversion? Merkle asks for anticipated satisfaction on the expected return over the next three months. He then matches responses with the best guess for the expected returns that survey participants provide. This might generate a bias, since returns are not exogenously given. Highly pessimistic participants provide subjective ratings on losses, while more optimistic or realistic participants provide ratings on gains.

Projection bias in satisfaction

Experiment I delivers no consistent pattern in terms of differences between anticipated and realized satisfaction. Especially for the loss domain, I cannot identify a robust loss aversion. But this could be due to the use of multiple treatments and resulting subgroups with as few as one participant. The experiment served as a first orientation for improvement. A simple asset with two return outcomes does not allow an estimation of the risk aversion coefficient, hence I focused on a single complex risky asset in the following experiments.

Experiment II shows a small but significant difference of 3.6 points between realized and anticipated satisfaction in the loss domain, and no significant projection bias in the gain domain. Still the projection bias in satisfaction does not induce a projection bias, i.e. an overestimation, in loss aversion. The estimated coefficient on anticipated satisfaction is even smaller than the coefficient for realized satisfaction. Experiment III resulted in almost identical averages for anticipated and realized satisfaction over all returns of the six possible, except of the largest return of 30% for which I report a statistically significant but small projection bias of 4 points.

The simple return distribution with six possible outcomes of equal probability might be the reason why participants were able to give similar statements for anticipated and realized returns. It is entirely possible that participants tried to perform well in remembering the "correct" level of satisfaction instead of giving truthful statements of emotion.

Why are these findings not in line with Kermer et al. (2006)? The authors record the happiness of participants after observing or actively playing a multi-round card gamble (and losing $4 of a $5 endowment). The group of participants forecasting their happiness after losing $4 in the game who observed a simulation anticipated a significantly lower happiness than those participants who actively played the forty round game themselves. Kermer et al. (2006) report that satisfaction with gains was overestimated and satisfaction with losses underestimated, hence they conclude that loss aversion itself is a forecasting error. A strong caveat relating to the analysis in Kermer et al. (2006) is that satisfaction can be influenced simply by the level of involvement. For example, Merkle et al. (2015) find that active involvement improves investor happiness. Given that personal responsibility changes the experience of outcomes (Zeelenberg et al. (1998)), it might be that experienced losses did not cause severe distress because participants were strongly engaged and felt they had done their best after 40 possibly entertaining rounds of a simple card game. Kermer et al. (2006) compare two groups with very different roles and very different levels of engagement and label their responses as anticipated and realized happiness, their results are therefore likely driven by other factors than a projection bias in predicting happiness.

Experience sampling

The effect of providing the information on the risky asset's return distribution in the form of experience sampling (Kaufmann and Weber (2013)) on the projection bias is assessed in experiment II. I can only report a significant decrease in anticipated satisfaction for the experience sampling group. The diminishing effect on the projection bias in satisfaction for the loss domain is very small and hence not statistically significant. It is possible that eliciting satisfaction on anticipated returns is a form of experience sampling that diminishes the effect of the information treatment. However, I only use the first anticipated satisfaction statement which is not affected by anchoring or an experience effect. If the elicitation had an effect on realized satisfaction, I should see a stronger overall projection bias in satisfaction.

Investment behavior and satisfaction

In experiments II and III, I analyze how investment decisions and satisfaction with investment outcomes are related. I find that a successful (failed) investment strategy increases (decreases) satisfaction with realized return outcomes only if participants are provided with the information on their initial allocation at the time of the elicitation of the realized satisfaction. Surprisingly, success or failure in the past investment round, i.e. higher vs lower allocation in the gain domain or lower vs higher allocation in the loss domain, does not influence future investment decisions.

A projection bias in satisfaction, i.e. a surprise in terms of higher realized than anticipated satisfaction, significantly affects future investment behavior only after outcomes in the gain domain, but coefficient estimates between experiment II and III reverse their sign and are therefore inconclusive. For investors that received an outcome in the loss domain, differences in satisfaction, i.e. learning that losses loom differently in reality compared to their anticipation, does not affect future investment decisions. The difference between experiments II and III, in communicating and reminding participants of their initial investment decision (experiment II) suggests that people base their decisions only on factors salient to them. It is hence not surprising that the allocation does not change between the first and the second round in experiments II and III, even after participants experience a much different satisfaction with an outcome than they expected. They might not even be aware of a difference, or not able to incorporate the surprise into their decision making process.

Conclusion

The existence of a projection bias in loss aversion and satisfaction on returns remains an open question. While the results of Kermer et al. (2006) can be explained by other factors, the findings of Merkle (2020) are backed by an extensive amount of robustness checks. Nevertheless, I find several factors that might critically bias his estimates. A perfect measure of realized satisfaction is difficult to obtain: In contrast to Merkle (2020)), I measure the realized satisfaction directly after return realization. The experience of winning or losing is thus fresh and salient to participants. Merkle (2020), on the other hand, uses quarterly surveys to elicit satisfaction. It is thus not clear how strongly coping mechanisms or other factors influence the happiness measures for past returns, resulting in "remembered utility" being measured rather than "instant utility" (Kahneman et al. (1997)). Merkle dismisses the possible effect of paper losses (see Imas (2016)), but it seems implausible that the quarterly returns of a portfolio are considered terminal by investors. Unless investors sell large parts of their assets, the portfolio returns of well diversified portfolios will be determined predominantly by market fluctuations, and it is hard to believe that losses are internalized on a repeated short-term basis given that equity investments have a positive long term expected return.

Furthermore, selection on the satisfaction and return data tuples might further bias Merkle's analyses. It seems that subjective ratings for expected returns and perceived returns deliver only few observations in the loss domain. The 5th percentile lies at -5% for expected returns, at -30% for perceived past returns, but at -52.1% for actual returns, whereas the 95th percentile only varies in a range of 5% between the return measures. Additionally, Merkle states that "participants clearly overestimate their past returns". If participants overestimate their past returns, they might also be less disappointed than they would have been, had they realized the actual (bad) performance, thus biasing the results. Research suggests that people have a hard time even stating their past portfolio performance (Glaser and Weber (2007)). It is therefore unclear whether and in which direction uncertainty about perceived returns could influence subjective happiness ratings. Finally, an additional selection bias arises because loss aversion is strongly correlated with risk taking. Observed outcomes are thus the result of preferences and previous decisions, and the success or failure of previous investment strategies will likely influence realized satisfaction.

Future research faces at least three challenges: First, eliciting subjective ratings for anticipated and realized returns using identical questions within the same survey while avoiding an anchoring effect. In the case of my third experiment, clever economics students possibly extrapolated over the six possible, equidistant returns of the risky asset. They might have stated a subjective rating that they thought to be appropriate, for example starting with zero satisfaction at the worst outcome and increasing linearly over the better outcomes. In that case, they might have focused on calculating the correct satisfaction for realized returns rather than relying on their true feelings and emotions. Second, satisfaction or happiness in relation to past events changes rapidly. For example, stubbing your little toe probably gives an infinitely negative happiness boost in instant utility, but five minutes later this impact has vanished. Ideally, we should monitor investor happiness directly, and for an extended period of time, after a return realization. But surveys and online experiments with numerically stated subjective ratings of happiness will not be able to achieve this. Repeated numerical self-assessments will be biased after the first round, which brings me back to the first issue. Third, investor happiness depends on previous choices and decisions and how successful these strategies are. It is even likely that this context influences the rating on anticipated returns differently than realized returns.

A fruitful next step could thus involve an additional experiment with regular people and a realistic complex return structure on the risky asset. A setup similar to experiment III should be chosen to maintain maximal independence of satisfaction statements and investment decisions. If a projection bias in satisfaction with observed investment losses is still absent, loss aversion might not be the result of a projection bias after all.

Bibliography

Abdellaoui, Mohammed, Han Bleichrodt, and Olivier L'Haridon. 2008. "A tractable method to measure utility and loss aversion under prospect theory". *Journal of Risk and Uncertainty* 36 (3): 245–266.

Abdellaoui, Mohammed, Han Bleichrodt, and Corina Paraschiv. 2007. "Loss Aversion Under Prospect Theory: A Parameter-Free Measurement". *Management Science* 53 (10): 1659–1674.

Acland, Dan, and Matthew R. Levy. 2015. "Naiveté, Projection Bias, and Habit Formation in Gym Attendance". *Management Science* 61 (1): 146–160.

Benartzi, S., and R. H. Thaler. 1995. "Myopic Loss Aversion and the Equity Premium Puzzle". *The Quarterly Journal of Economics* 110 (1): 73–92.

Booij, Adam S., and Gijs van de Kuilen. 2009. "A parameter-free analysis of the utility of money for the general population under prospect theory". *Journal of Economic Psychology* 30 (4): 651–666.

Booij, Adam S., van Praag, Bernard M. S., and Gijs van de Kuilen. 2010. "A parametric analysis of prospect theory's functionals for the general population". *Theory and Decision* 68 (1): 115–148.

Bradbury, Meike, Thorsten Hens, and Stefan Zeisberger. 2019. "How persistent are the effects of experience sampling on investor behavior?" *Journal of Banking and Finance.*

Busse, Meghan R., Devin G. Pope, Jaren C. Pope, and Jorge Silva-Risso. 2012. "Projection Bias in the Car and Housing Markets". *NBER Working Paper,* number 18212.

Cao, Melanie, and Jason Wei. 2002. "Stock market returns: A temperature anomaly". *Working Paper; SSRN.com.*

Conlin, Michael, Ted O'Donoghue, and Timothy J. Vogelsang. 2007. "Projection Bias in Catalog Orders". *American Economic Review* 97 (4): 1217–1249.

Dowling, Michael M., and Brian M. Lucey. 2003. "The Role of Feelings in Investor Decision-Making". *SSRN Electronic Journal.*

Glaser, Markus, and Martin Weber. 2007. "Why inexperienced investors do not learn: They do not know their past portfolio performance". *Finance Research Letters* 4 (4): 203–216.

Goetzmann, William N., and Ning Zhu. 2005. "Rain or Shine: Where is the Weather Effect?" *European Financial Management* 11 (5): 559–578.

Grosshans, Daniel, and Stefan Zeisberger. 2018. "All's well that ends well? On the importance of how returns are achieved". *Journal of Banking & Finance* 87:397–410.

Haliassos, Michael, and Carol C. Bertaut. 1995. "Why Do So Few Hold Stocks?" *Economic Journal* 105 (432): 1110–1129.

Hirshleifer, David, and Siew Hong Teoh. 2003. "Limited attention, information disclosure, and financial reporting". *Journal of Accounting and Economics* 36 (1-3): 337–386.

Imas, Alex. 2016. "The Realization Effect: Risk-Taking After Realized Versus Paper Losses". *American Economic Review* 106 (8): 2086–2109.

Imas, Alex, Sally Sadoff, and Anya Samek. 2016. "Do People Anticipate Loss Aversion?" *Management Science* 63 (5): 1271–1656.

Kahneman, Daniel, and Jackie Snell. 1990. "Predicting utility". In *Insights in decision making: A tribute to Hillel J. Einhorn,* 295–310. Chicago, IL, US: University of Chicago Press.

———. 1992. "Predicting a changing taste: Do people know what they will like?" *Journal of Behavioral Decision Making* 5 (3): 187–200.

Kahneman, Daniel, and Amos Tversky. 1979. "Prospect Theory: An Analysis of Decision under Risk". *Econometrica* 47 (2): 263.

———. 1984. "Choices, values, and frames". *American Psychologist* 39 (4): 341–350.

Kahneman, Daniel, Peter P. Wakker, and Rakesh Sarin. 1997. "Back to Bentham? Explorations of Experienced Utility". *The Quarterly Journal of Economics* 112 (2): 375–405.

Kamstra, Mark J., Lisa A. Kramer, and Maurice D. Levi. 2003. "Winter Blues: A SAD Stock Market Cycle". *American Economic Review* 93 (1): 324–343.

Kaufmann, Christine, and Martin Weber. 2013. "Sometimes less is more – The influence of information aggregation on investment decisions". *Journal of Economic Behavior & Organization* 95:20–33.

Kermer, Deborah A., Erin Driver-Linn, Timothy D. Wilson, and Daniel T. Gilbert. 2006. "Loss aversion is an affective forecasting error". *Psychological science* 17 (8): 649–653.

Kőszegi, Botond, and Matthew Rabin. 2008. "Choices, situations, and happiness". *Journal of Public Economics* 92 (8-9): 1821–1832.

Loewenstein, George. 2005. "Projection bias in medical decision making". *Medical decision making : an international journal of the Society for Medical Decision Making* 25 (1): 96–105.

Loewenstein, George, Ted O'Donoghue, and Matthew Rabin. 2003. "Projection Bias in Predicting Future Utility". *The Quarterly Journal of Economics* 118 (4): 1209–1248.

Loewenstein, George, Elke U. Weber, Christopher K. Hsee, and Ned Welch. 2001. "Risk as feelings". *Psychological Bulletin* 127 (2): 267–286.

Loomes, Graham, and Robert Sugden. 1982. "Regret Theory: An Alternative Theory of Rational Choice Under Uncertainty". *The Economic Journal* 92 (368): 805.

Mankiw, N.Gregory, and Stephen P. Zeldes. 1991. "The consumption of stockholders and nonstockholders". *Journal of Financial Economics* 29 (1): 97–112.

Mehra, Rajnish, and Edward C. Prescott. 1985. "The equity premium: A puzzle". *Journal of Monetary Economics* 15 (2): 145–161.

Merkle, Christoph. 2020. "Financial Loss Aversion Illusion". *Review of Finance* 24 (2): 381–413.

Merkle, Christoph, Daniel P. Egan, and Greg B. Davies. 2015. "Investor happiness". *Journal of Economic Psychology* 49:167–186.

Merton, Robert C. 1969. "Lifetime Portfolio Selection under Uncertainty: The Continuous-Time Case". *The Review of Economics and Statistics* 51 (3): p 247–257.

Samuelson, Paul A. 1969. "Lifetime Portfolio Selection by Dynamic Stochastic Programming". *The Review of Economics and Statistics* 51 (3): 239–246.

Saunders, Edward M, Jr. 1993. "Stock Prices and Wall Street Weather". *American Economic Review* 83 (5): 1337–1345.

Schmittmann, Jochen M., Jenny Pirschel, Steffen Meyer, and Andreas Hackethal. 2015. "The Impact of Weather on German Retail Investors*". *Review of Finance* 19 (3): 1143–1183.

Shefrin, Hersh, and Meir Statman. 1985. "The Disposition to Sell Winners Too Early and Ride Losers Too Long: Theory and Evidence". *The Journal of Finance* 40 (3): 777–790.

Shiv, Baba, George Loewenstein, and Antoine Bechara. 2005. "The dark side of emotion in decision-making: when individuals with decreased emotional reactions make more advantageous decisions". *Brain research. Cognitive brain research* 23 (1): 85–92.

Shiv, Baba, George Loewenstein, Antoine Bechara, Hanna Damasio, and Antonio R. Damasio. 2005. "Investment Behavior and the Negative Side of Emotion". *Psychological science* 16 (6): 435–439.

Strahilevitz, Michal Ann, Terrance Odean, and Brad M. Barber. 2011. "Once Burned, Twice Shy: How Naive Learning, Counterfactuals, and Regret Affect the Repurchase of Stocks Previously Sold". *Journal of Marketing Research* 48 (SPL): S102–S120.

Thaler, Richard H. 1990. "Anomalies: Saving, Fungibility, and Mental Accounts". *Journal of Economic Perspectives* 4 (1): 193–205.

Tversky, Amos, and Daniel Kahneman. 1992. "Advances in prospect theory: Cumulative representation of uncertainty". *Journal of Risk and Uncertainty* 5 (4): 297–323.

Wilson, Timothy D., and Daniel T. Gilbert. 2003. "Affective Forecasting". *Advances in Experimental Social Psychology,* number 35: 345–411.

Zeelenberg, Marcel, Wilco W. van Dijk, and Antony S.R Manstead. 1998. "Reconsidering the Relation between Regret and Responsibility". *Organizational Behavior and Human Decision Processes* 74 (3): 254–272.

Zeisberger, Stefan, Dennis Vrecko, and Thomas Langer. 2012. "Measuring the time stability of Prospect Theory preferences". *Theory and Decision* 72 (3): 359–386.

Appendix A - Overview of Experiments

Table A.1.
Experiment overview - participants

Experiment	I	II	III
Investment task	Allocate €1000	Allocate US$1000	Allocate €1000
Information treatment	Experience sampling or Description of return distribution	Experience sampling or Description of return distribution	Description of return distribution
Risky Asset(s)	Simple (2 outcomes) or Complex (continuous)	Complex (continuous)	Die (6 outcomes)
Participants	Econ students	M-turk	Econ students
N	318	427	402
Age	-	35.7	-
Female (%)	36.2	36.2	40.8
Own mutual funds (%)	9.1	54.9	14.7
Interested in stocks (%)	73.6	75.9	-
Average time (minutes)	≈ 25 (20 part 1, 5 part 2)	15.8	28.6 (23.1 part 1, 25.5 part 2)
Average allocation			
initial (anticipated)	€580.50	$582.64	€619.30
2nd round (realized)	€586.10	$574.73	€625
Final outcome	€1045.30	$1023.61	€1035.47
Average Payout	€10.45 (calculated)	$10.75 (calculated)	€10.35

Table A.2.
Experiment overview - results

Experiment	I	II	III
Loss aversion coeff >1?			
Anticipated	No	No	Yes (λ_{Antic} = 1.138, p-val=0.042)
Realized	No	No	No
Projection bias			
in loss aversion?	No	No $\lambda_{Real} > \lambda_{Antic}$	Yes
in satisfaction?			
Gains	Yes (only complex asset, p-val< 10%)	No	No
Losses	No	Yes (significant t-test on differences and regression coeff.)	No
Impact of experience sampling on satisfaction and projection bias?	Smaller satisfaction on realized gains for exp. sampling (*)	Smaller satisfaction on anticipated gains for exp. sampling (**); / / Proj. bias estimated larger in gain domain, slightly smaller in loss domain both insignificant	-
Impact of initial allocation on realized satisfaction and projection bias?	-	Gains: higher allocation -> higher satisf. Losses: higher allocation -> lower satisf.	No
Impact of projection bias in satisfaction on investment (re-)allocation decision?	-	negative for gain domain, insignificant for losses	No

Table A.3.
Variable description

Satisfaction variables	
Anticipated satisfaction	Satisfaction anticipated for potential returns: Integer betwen 0 (extremely dissatisfied) and 100 (extremely satisfied), see experimental instructions Appendix C.
Realized satisfaction	Satisfaction on realized returns: Integer betwen 0 (extremely dissatisfied) and 100 (extremely satisfied), see experimental instructions Appendix C.
Difference in satisfaction	Realized - anticipated satisfaction, in experiment 2 and 3 the participant stated the anticipated satisfaction exactly on the realized return. For experiment 1 I use interpolation to calculate return specific differences in satisfaction.
Allocation variables	
Initial allocation	First investment into the risky asset: An integer value between 0 and 1000.
Second allocation	Second investment into the risky asset, after the first round return was realized: An integer value between 0 and 1000.
Difference in allocation	Second - initial allocation.
Control variables - participants	
Female	Equals 1 if the participant is female, and equals zero otherwise.
Risk tolerance	5-point Likert scale from 1 (not willing to accept any risk) to 5 (willing to accept substantial risk to potentially earn a greater return).
Percieved Risk	9-point Likert scale from 1 (not risky at all) to 9 (very risk).
Loss probability	In how many out of 100 cases would you expect the return to be negative (lower than 0%): Integer value between 0 an 100.
Interested in stocks	Equals 1 if the participant stated to be interested in stock investments, and equals zero otherwise.
Owns mutual funds	Equals 1 if the participant stated to own mutual funds, and equals zero otherwise.
Financial literacy	The sum of individual question answered correctly: Score between 0 and 12 (experiment 1), 0 and 4 (experiment 2 and 3). Experiment 1 includes an additional variable that equals 1 if the participant answered the diversification question correctly, and equals zero otherwise.
Numeracy	The sum of individual question answered correctly: Score between 0 and 4 (only in experiment 1 and 2).
Comprehension	The sum of individual question answered correctly: Score between 0 and 8 (experiment 2), or 0 and 2 (experiment 3)

Table A.3.
(Continued)

Control variables - return related	
Return	Return drawn for the risky asset, if not stated otherwise it is the realized return after investment.
Losses	Second investment into the risky asset, after the first round return was realized. An integer value between 0 and 1000.
Gains	Equals 1 if the observed return is positive, and equals zero otherwise.
Anticipated	Equals 1 if satisfaction is measured on anticipated returns, and equals zero otherwise.
Realized	Equals 1 if satsifaction is measured on realized returns, and equals zero otherwise.
Control variables - treaments	
Description	Equals 1 if participant received the descriptive information treatment, and equals zero otherwise.
Experience	Equals 1 if participant received the experience sampling information treatment, and equals zero otherwise.
Simple	Equals 1 if participant received the simple risky asset, and equals zero otherwise.
Complex	Equals 1 if participant received the complex risky asset, and equals zero otherwise.

Appendix B - Experimental Results

Table B.1.
Projection bias in loss aversion - summary of regression tests
This table reports OLS regression estimates for the calculation of loss aversion coefficients. The dependent variable is satisfaction. For each participant I have two observations that share the same realized return, one for anticipated satisfaction, and one for realized satisfaction. I regress satisfaction on the interactions of return and a dummy for the loss domain (row 1) and of return and a dummy for the gain domain (row 2). The loss aversion coefficient is equal to the coefficient in row 1 divided by the coefficient in row 2. To test for differences in loss aversion between anticipated returns and realized returns, I additionally interact the return interactions with a dummy for anticipated returns (row 3 and 4). Other independent variables include the dummy for anticipated returns without interactions, and all participant control variables as specified in table A.3. The loss aversion coefficient for anticipated returns, λ_{Antic}, is equal to $(\beta_1 + \beta_3)/(\beta_2 + \beta_4)$ (subscripts indicate the row for each coefficient estimate in the table below). The loss aversion coefficient for realized returns, λ_{Real}, is equal to $(\beta_1)/(\beta_2)$. I report the p-values on Wald-tests to test if the loss aversion coefficients are different from 1 (loss neutrality) and different from each other. Columns represent separate regressions for each experiment. P-values on coefficients are provided in parentheses and are based on heteroscedasticity robust z-statistics. ***, **, and * denote statistical signifficance at the 1%, 5%, and 10% levels, respectively.

	(1) Experiment I	(2) Experiment II	(3) Experiment III
Return·Losses	1.569***	1.257***	1.769***
	(0.000)	(0.000)	(0.000)
Return·Gains	1.331***	1.122***	1.759***
	(0.000)	(0.000)	(0.000)
Return·Losses·Anticipated	0.023	−0.073	0.088
	(0.926)	(0.238)	(0.258)
Return·Gains·Anticipated	0.059	0.188**	−0.127***
	(0.750)	(0.017)	(0.008)
Anticipated	−2.197	−4.426***	1.521
	(0.519)	(0.003)	(0.105)
Constant	44.064***	59.878***	44.745***
	(0.000)	(0.000)	(0.000)
Participant controls	✓	✓	✓
R-Squared	0.68	0.62	0.83
Observations	431	824	804
Loss aversion coefficient estimates and Wald-tests			
λ_*Antic*	1.145	0.904	1.138
P-value (λ_*Antic* = 1)	(0.512)	(0.574)	(0.042)
λ_*Real*	1.178	1.121	1.005
P-value (λ_*Real* = 1)	(0.434)	(0.584)	(0.926)
P-value (λ_*Antic* = λ_*Real*)	(0.914)	(0.062)	(0.031)

B.I. Tables Experiment I

Table B.2.
Anticipated and realized satisfaction - averaged by return ranges (Experiment 1)
This table reports the stated satisfaction in anticipation or on realized returns. Satisfaction is averaged over ranges of returns represented in columns. The first column shows averages for returns smaller or equal -20%, the second column shows averages for returns greater than -20% and smaller or equal -10% and so on. The last two columns show averages for the loss and gain domain respectively. The first set of rows reports the number of observation in each return range (columns) and sample split by treatment (rows). For the simple asset returns are discrete, either -10% or +20%. The anticipated satisfaction is interpolated (extrapolated for the tails of the return distribution) for the realized return to match the corresponding realized satisfaction. The last set of rows are the differences in satisfaction (realized-anticipated). T-tests results (two-sided) are provided as follows: ***, **, and * denote statistical significance of differences at the 1%, 5%, and 10% levels, respectively.

	<= −20%	<= −10%	<= 0%	<= 10%	<= 20%	<= 30%	> 30%	<= 0%	> 0%
Number of observations									
Overall	10	73	31	54	97	36	17	114	204
Simple Asset		50			52			50	52
Description		26			24			26	24
Experience		24			28			24	28
Complex Asset	10	23	31	54	45	36	17	64	152
Description	1	14	16	31	19	18	10	31	78
Experience	9	9	15	23	26	18	7	33	74
Anticipated sastisfaction									
Overall	10.4	24.2	32.8	56.6	75.9	85.5	84.6	25.3	73.2
Simple Asset		26.7			79.2			26.7	79.2
Description		29.3			79.3			29.3	79.3
Experience		23.8			79.2			23.8	79.2
Complex Asset	10.4	18.7	32.8	56.6	72.0	85.5	84.6	24.3	71.1
Description	0.0	15.8	33.2	57.4	71.6	83.4	84.5	24.3	70.3
Experience	11.5	23.4	32.5	55.5	72.3	87.6	84.7	24.3	72.0
Realized satisfaction									
Overall	16.6	23.1	25.9	59.8	76.9	85.8	85.8	23.3	74.7
Simple Asset		23.5			78.1			23.5	78.1
Description		23.0			83.0			23.0	83.0
Experience		24.0			73.9			24.0	73.9
Complex Asset	16.6	22.2	25.9	59.8	75.5	85.8	85.8	23.1	73.5
Description	7.0	25.2	23.9	59.0	76.0	86.7	87.9	23.9	73.2
Experience	17.7	17.4	28.1	60.9	75.2	84.8	82.7	22.3	73.8
Realized - Anticipated									
Overall	6.2	−1.1	-6.9^{**}	3.2	1.0	0.3	1.2	−2.1	1.5
Simple Asset		−3.2			−1.1			−3.2	−1.1
Description		−6.3			3.7			−6.3	3.7
Experience		0.2			−5.3			0.2	−5.3
Complex Asset	6.2	3.4	-6.9^{**}	3.2	3.5^{*}	0.3	1.2	−1.2	2.4^{*}
Description	7.0^{***}	9.4	-9.3^{**}	1.6	4.4	3.3	3.4	−0.3	2.9
Experience	6.1	−5.9	−4.4	5.4	2.9	−2.7	−2.0	−1.9	1.8

Table B.3.
Experience sampling and return complexity's impact on satisfaction - regression tests
This table reports estimates of a OLS regression analyses to test if experience sampling and the risky asset's complexity influence investors satisfaction with returns and thus have an impact on a potential projection bias in satisfaction. The dependent variable is anticipated satisfaction for models 1 and 2, and realized satisfaction for models 3 and 4. I split the participants by the realized return. Models 1 and 3, include only participants with positive returns (gains) and models 2 and 4 include participants with negative or zero returns. For each participant I include only one observation per model. Anticipated satisfaction includes only the observation per participant on the return that was drawn (realized) for the investment into the risky asset and is therefore interpolated. For the variables of main interest, I regress satisfaction on a dummy for the experience information treatment, a dummy for the complex asset and the interaction of the two. Control variables include the return corresponding to the stated satisfaction, and all participant control variables as specified in table A.3. P-values on coefficients are provided in parentheses and are based on heteroscedasticity robust z-statistics. ***, **, and * denote statistical signifficance at the 1%, 5%, and 10% levels, respectively.

	(1)	(2)	(3)	(4)
Dep.Var.: Satisfaction	Anticipated	Anticipated	Realized	Realized
Return domain	Gains	Losses	Gains	Losses
Experience	0.107	−5.685	−8.364*	4.302
	(0.983)	(0.249)	(0.060)	(0.227)
Complex	−3.461	−6.123	−4.764	7.470
	(0.443)	(0.216)	(0.223)	(0.114)
Complex·Experience	1.011	10.766	7.188	−2.903
	(0.851)	(0.105)	(0.161)	(0.605)
Return	1.009***	1.210***	1.001***	0.520**
	(0.000)	(0.000)	(0.000)	(0.027)
Female	0.867	1.475	2.730	9.427***
	(0.754)	(0.667)	(0.253)	(0.001)
Risk tolerance	−1.323	1.720	2.480	4.283**
	(0.383)	(0.481)	(0.150)	(0.024)
Perceived risk	−0.517	1.222	−0.218	0.933
	(0.499)	(0.224)	(0.754)	(0.297)
Owns mutual funds	−5.147	14.336	−5.196	39.926***
	(0.142)	(0.236)	(0.251)	(0.001)
Interested in stocks	1.252	0.528	2.436	−1.999
	(0.632)	(0.894)	(0.341)	(0.622)
Loss probability	0.111	−0.035	−0.019	0.263**
	(0.283)	(0.796)	(0.826)	(0.027)
Financial literacy score	0.516	−1.289	0.074	−1.797*
	(0.501)	(0.306)	(0.902)	(0.067)
Fin.lit. diversification	1.984	6.923	0.983	1.178
	(0.624)	(0.161)	(0.736)	(0.811)
Financial numeracy score	1.344	−0.168	0.274	0.708
	(0.205)	(0.893)	(0.787)	(0.557)
R-Squared	0.357	0.206	0.347	0.364
Observations	204	114	204	114

B.II. Tables Experiment II

Table B.4.
Anticipated and Experienced Satisfaction

This table reports the stated satisfaction in anticipation or on realized returns. Satisfaction is averaged over ranges of returns represented in columns. The first column shows averages for returns smaller or equal -20%, the second column shows averages for returns greater than -20% and smaller or equal -10% and so on. The last two columns show averages for the loss and gain domain respectively. Rows represent sample splits by information treatment. The first set of rows reports the number of observation in each return range (columns). The last set of rows show the differences in satisfaction (realized-anticipated). T-tests results (two-sided) are provided as follows: ***, **, and * denote statistical significance of differences at the 1%, 5%, and 10% levels, respectively.

	<= −20%	<= −10%	<= 0%	<= 10%	<= 20%	<= 30%	> 30%	<= 0%	> 0%
Number of observations									
Overall	59	52	74	68	67	58	47	185	240
Description	29	26	39	34	33	29	23	94	119
Experience	30	26	35	34	34	29	24	91	121
Anticipated sastisfaction									
Overall	12.8	11.3	26.0	56.2	76.4	83.5	90.4	17.7	75.3
Description	10.3	12.7	29.0	57.4	82.8	86.0	91.7	18.7	78.1
Experience	15.2	9.9	22.7	55.0	70.2	80.9	89.3	16.6	72.6
Realized satisfaction									
Overall	14.5	14.7	31.4	59.9	75.6	83.8	87.7	21.3	75.5
Description	11.3	17.6	34.3	62.1	77.2	85.4	91.7	22.6	77.7
Experience	17.5	11.7	28.2	57.7	74.0	82.1	83.8	20.0	73.3
Realized - Anticipated									
Overall	1.7	3.3	5.4^{**}	3.7^{**}	−0.8	0.3	−3.0	3.6^{***}	0.3
Description	1.0	4.9	5.3	4.7^{*}	-5.6^{*}	−0.6	0.0	3.9^{**}	−0.3
Experience	2.4	1.8	5.5^{**}	2.7	3.8	1.2	−5.8	3.4^{*}	1.0

Table B.5.
Experience sampling's impact on satisfaction - regression tests (Exp.II)
This table reports estimates of OLS regression analyses to test if experience sampling influence investors' satisfaction with returns and thus have an impact on a potential projection bias in satisfaction. The dependent variables are anticipated satisfaction for model 1, realized satisfaction for model 2, and the difference between realized - anticipated satisfaction (the projection bias) for models 3 and 4. For each participant I include only one observation per model. Anticipated satisfaction includes only the observation per participant on the return that was drawn (realized) for the investment into the risky asset. For the variables of main interest, I regress satisfaction on a dummy for the experience information treatment, a dummy for the loss domain and their interaction. Control variables include the return corresponding to the stated satisfaction, its interaction with the loss domain dummy, and all participant control variables as specified in table A.3. P-values on coefficients are provided in parentheses and are based on heteroscedasticity robust z-statistics. ***, **, and * denote statistical signifficance at the 1%, 5%, and 10% levels, respectively.

Dep.Var.: Satisfaction	(1) Anticipated	(2) Realized	(3) Difference	(4) Difference
Experience	−5.897**	−3.554	2.351	2.343
	(0.015)	(0.141)	(0.265)	(0.263)
Losses	−36.561***	−32.575***	5.050**	3.986
	(0.000)	(0.000)	(0.041)	(0.286)
Experience•Losses	3.649	0.614	−3.122	−3.035
	(0.347)	(0.886)	(0.365)	(0.374)
Return	0.782***	0.628***		−0.155
	(0.000)	(0.000)		(0.124)
Return•Losses	−0.344*	−0.057		0.287**
	(0.057)	(0.762)		(0.025)
Age	−0.145*	−0.178**	−0.018	−0.033
	(0.050)	(0.030)	(0.793)	(0.632)
Female	0.308	2.023	1.354	1.715
	(0.886)	(0.359)	(0.455)	(0.351)
Risk tolerance	−1.591	−2.382	−0.761	−0.790
	(0.173)	(0.112)	(0.540)	(0.529)
Perceived risk	−0.190	−0.010	0.182	0.180
	(0.758)	(0.987)	(0.711)	(0.713)
Financial literacy score	−0.642	−0.216	0.551	0.427
	(0.547)	(0.845)	(0.456)	(0.568)
Financial numeracy score	1.380*	1.959***	0.500	0.579
	(0.053)	(0.009)	(0.419)	(0.358)
Comprehension score	−0.035	0.471	0.482	0.506
	(0.969)	(0.613)	(0.495)	(0.482)
Constant	73.213***	71.720***	−5.207	−1.493
	(0.000)	(0.000)	(0.436)	(0.828)
R-Squared	0.724	0.662	0.025	0.038
Observations	412	412	412	412

Table B.6.
Projection bias in satisfaction - regression tests (Exp.II)
This table reports estimates of OLS regression analyses to test if there is a projection bias in satisfaction, i.e. a positive difference between realized-anticipated satisfaction, especially in the loss domain. The dependent variables are the difference between realized - anticipated satisfaction (the projection bias) for models 1 to 3, anticipated satisfaction for model 4 and realized satisfaction for model 5. The main variables of interest are the constant, a dummy for the loss domain, and the allocation to the risky asset prior to the draw of the realized return (initial allocation). Control variables include a dummy for experience sampling and all participant control variables as specified in table A.3. P-values on coefficients are provided in parentheses and are based on heteroscedasticity robust z-statistics. ***, **, and * denote statistical signifficance at the 1%, 5%, and 10% levels, respectively.

	(1)	(2)	(3)	(4)	(5)
Dep.Var.: Satisfaction	Difference	Difference	Difference	Anticipated	Realized
Losses		3.312**	14.294***	−55.293***	−40.999***
		(0.045)	(0.008)	(0.000)	(0.000)
Initial allocation			0.009*	0.005	0.014**
			(0.070)	(0.373)	(0.011)
Init. allocation·Losses			−0.019**	−0.004	−0.023**
			(0.017)	(0.562)	(0.011)
Experience			1.072	−5.044**	−3.972*
			(0.542)	(0.027)	(0.096)
Age			0.016	−0.183**	−0.167*
			(0.811)	(0.031)	(0.056)
Female			1.160	1.155	2.315
			(0.512)	(0.618)	(0.322)
Risk tolerance			−0.426	−1.473	−1.898
			(0.677)	(0.290)	(0.208)
Perceived risk			0.322	−0.056	0.266
			(0.483)	(0.935)	(0.675)
Financial literacy score			0.118	−1.280	−1.162
			(0.878)	(0.291)	(0.346)
Financial numeracy score			0.788	1.696**	2.485***
			(0.213)	(0.049)	(0.004)
Comprehension score			0.321	0.161	0.482
			(0.646)	(0.859)	(0.607)
Constant	1.762**	0.321	−10.962	85.734***	74.772***
	(0.030)	(0.753)	(0.122)	(0.000)	(0.000)
R-Squared	0.000	0.010	0.049	0.661	0.619
Observations	425	425	412	412	412

Table B.7.
Second - Initial allocation (risk re-allocation) and the Projection Bias (Exp.II)
This table reports estimates of OLS regression analyses to test the impact of satisfaction on the investment behavior. More specific, if the difference between realized-anticipated satisfaction influences risk taking after return realization. The dependent variables are the difference between second - initial allocations for models 1 to 4, and the second allocation for model 5. The main variables of interest are the difference in satisfaction (the projection bias) and its interaction with a dummy for the loss domain. Control variables include a dummy for the loss domain, the allocation to the risky asset prior to the draw of the realized return (initial allocation), and and all participant control variables as specified in table A.3. P-values on coefficients are provided in parentheses and are based on heteroscedasticity robust z-statistics. ***, **, and * denote statistical signifficance at the 1%, 5%, and 10% levels, respectively.

Dep.Var.: Allocation	(1) Difference	(2) Difference	(3) Difference	(4) Difference	(5) Realized
Satisfaction difference		−2.38*	−2.46*	−1.56	−1.56
		(0.092)	(0.081)	(0.213)	(0.213)
Satisf.difference•Losses		2.45	2.59	1.08	1.08
		(0.121)	(0.109)	(0.453)	(0.453)
Initial allocation				−0.34***	0.66***
				(0.000)	(0.000)
Init.allocation•Losses				0.02	0.02
				(0.784)	(0.784)
Losses	−206.65***	−207.68***	−204.97***	−208.84***	−208.84***
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Age			−2.38*	−2.79**	−2.79**
			(0.069)	(0.020)	(0.020)
Female			41.96	39.94	39.94
			(0.127)	(0.120)	(0.120)
Risk tolerance			18.89	50.03***	50.03***
			(0.249)	(0.002)	(0.002)
Perceived risk			14.89*	6.34	6.34
			(0.064)	(0.399)	(0.399)
Financial literacy score			−4.72	0.60	0.60
			(0.761)	(0.965)	(0.965)
Financial numeracy score			5.74	6.53	6.53
			(0.573)	(0.498)	(0.498)
Comprehension score			−10.44	−9.97	−9.97
			(0.284)	(0.265)	(0.265)
Constant	81.08***	81.85***	69.87	207.64**	207.64**
	(0.000)	(0.000)	(0.447)	(0.028)	(0.028)
Effect for Losses		0.072	0.124	−0.483	−0.483
P-value		(0.919)	(0.875)	(0.490)	(0.490)
R-squared	0.131	0.141	0.173	0.281	0.493
Observations	425	425	412	412	412

B.III. Tables Experiment III

Table B.8.
Anticipated and Experienced Satisfaction
This table reports the stated satisfaction in anticipation or on realized returns. Satisfaction is averaged over ranges of returns represented in columns. The first column shows averages for returns smaller or equal -20%, the second column shows averages for returns greater than -20% and smaller or equal -10% and so on. The last two columns show averages for the loss and gain domain respectively. Rows represent sample splits by information treatment. The first reports the number of observation in each return range (columns). The last set of rows shows the differences in satisfaction (realized-anticipated). The rows named "1st Round" show the stated satisfaction on returns drawn after the first investment round, the rows names "2nd Round", accordingly, show the stated satisfaction on returns drawn after the second investment round. T-tests results (two-sided) are provided as follows: ***, **, and * denote statistical significance of differences at the 1%, 5%, and 10% levels, respectively.

	−20%	−10%	0%	10%	20%	30%	<= 0%	> 0%
Number of observations	71	68	47	61	82	73	186	216
Anticipated sastisfaction								
1st Round	11.1	24.6	38.0	67.0	85.7	89.8	22.9	81.8
2nd Round	9.9	24.1	45.7	66.3	80.9	93.6	26.0	80.3
Realized satisfaction								
1st Round	11.8	23.7	36.8	65.2	84.5	93.8	22.5	82.2
2nd Round	8.6	22.0	38.9	66.5	81.7	93.7	22.8	80.6
Realized - Anticipated								
1st Round	0.6	−0.9	−1.3	−1.8	−1.2	4.0***	−0.4	0.4
2nd Round	0.8	0.6	0.6	0.2	0.3	0.0	0.0	0.0

Table B.9.
Projection bias in satisfaction - regression tests (Exp.III)

This table reports estimates of OLS regression analyses to test if there is a projection bias in satisfaction, i.e. a positive difference between realized-anticipated satisfaction, especially in the loss domain. The dependent variables are the difference between realized - anticipated satisfaction (the projection bias) for models 1 to 3, anticipated satisfaction for model 4 and realized satisfaction for model 5. The main variables of interest are the constant, a dummy for the loss domain, and the allocation to the risky asset prior to the draw of the realized return (initial allocation). Control variables include all participant control variables as specified in table A.3. P-values on coefficients are provided in parentheses and are based on heteroscedasticity robust z-statistics. ***, **, and * denote statistical signifficance at the 1%, 5%, and 10% levels, respectively.

Dep.Var.: Satisfaction	(1) Difference	(2) Difference	(3) Difference	(4) Anticipated	(5) Realized
Losses	−0.195	−0.179	−1.785	−57.458***	−59.243***
	(0.892)	(0.902)	(0.677)	(0.000)	(0.000)
Initial allocation			0.001	0.002	0.003
			(0.802)	(0.748)	(0.649)
Init.allocation·Losses			0.003	0.002	0.005
			(0.717)	(0.805)	(0.584)
Female		0.785	0.824	−2.277	−1.453
		(0.540)	(0.536)	(0.291)	(0.513)
Risk tolerance		−0.224	−0.407	−3.091**	−3.498***
		(0.831)	(0.687)	(0.047)	(0.009)
Perceived risk		−0.126	−0.088	2.100***	2.012***
		(0.780)	(0.849)	(0.003)	(0.003)
Financial literacy score		−0.075	−0.049	0.730	0.681
		(0.940)	(0.961)	(0.668)	(0.671)
Comprehension score		−1.209	−1.364	1.290	−0.074
		(0.593)	(0.547)	(0.724)	(0.981)
Constant	0.087	3.583	3.666	68.623***	72.289***
	(0.908)	(0.548)	(0.565)	(0.000)	(0.000)
R-squared	0.000	0.002	0.004	0.636	0.639
Observations	402	402	402	402	402

Table B.10.
Second - Initial allocation (risk re-allocation) and the Projection Bias (Exp.III)
This table reports estimates of OLS regression analyses to test the impact of satisfaction on the investment behavior. More specific, if the difference between realized-anticipated satisfaction influences risk taking after return realization. The dependent variables are the difference between realized - initial allocations for models 1 to 4, and the second allocation for model 5. The main variables of interest are the difference in satisfaction (the projection bias) and its interaction with a dummy for the loss domain. Control variables include a dummy for the loss domain, the allocation to the risky asset prior to the draw of the realized return (initial allocation), and and all participant control variables as specified in table A.3. P-values on coefficients are provided in parentheses and are based on heteroscedasticity robust z-statistics. ***, **, and * denote statistical signifficance at the 1%, 5%, and 10% levels, respectively.

	(1)	(2)	(3)	(4)	(5)
Dep.Var.: Allocation	Difference	Difference	Difference	Difference	Realized
Losses	−21.90	−21.68	−20.43	−40.05	−40.05
	(0.439)	(0.443)	(0.474)	(0.636)	(0.636)
Satisfaction difference		2.18*	2.03*	2.34**	2.34**
		(0.075)	(0.088)	(0.031)	(0.031)
lossXsatisfDiff		−1.85	−1.53	−1.59	−1.59
		(0.324)	(0.411)	(0.353)	(0.353)
Initial allocation				−0.56***	0.44***
				(0.000)	(0.000)
Init.allocation•Losses				0.01	0.01
				(0.929)	(0.929)
Female			−6.93	−42.87*	−42.87*
			(0.791)	(0.059)	(0.059)
Risk tolerance			−15.94	31.44**	31.44**
			(0.267)	(0.042)	(0.042)
Perceived risk			4.56	−4.30	−4.30
			(0.564)	(0.545)	(0.545)
Financial literacy score			43.14**	31.62*	31.62*
			(0.029)	(0.055)	(0.055)
Comprehension score			−57.11	−6.09	−6.09
			(0.107)	(0.834)	(0.834)
Constant	13.27	13.08	48.07	239.84**	239.84**
	(0.366)	(0.371)	(0.636)	(0.011)	(0.011)
Effect for Losses		0.335	0.498	0.747	0.747
P-value		(0.813)	(0.727)	(0.571)	(0.571)
R-squared	0.002	0.009	0.030	0.277	0.268
Observations	402	402	402	402	402

Appendix C - Experimental Instructions

C.I. Running order of experiments

All experiments start with an introduction, that describes the process of the experiment, collect participant ids, and end with displaying the participants pay-off after investment and return draws.

Experiment I

1. Return information treatment (simple vs complex, experience vs description)
2. Riskiness and expectations
3. Investment decision I
4. Anticipated satisfaction (with display of risky allocation)
5. Financial literacy
6. Demographics
7. BREAK for second part
8. Return information treatment (review)
9. Draw of risky asset return, realized satisfaction
10. Investment decision II

Experiment II

1. Comprehension
2. Return information treatment (experience vs description)
3. Riskiness and expectations
4. Investment decision I
5. Anticipated satisfaction (with display of risky allocation)
6. Demographics
7. Numerical literacy
8. Draw of risky asset return, realized satisfaction
9. Investment decision II
10. Financial literacy

Experiment III

1. Baseline Happiness
2. Comprehension

3. Return information treatment (dice)
4. Anticipated satisfaction (no risky allocatino so far)
5. Riskiness and expectations
6. Investment decision I
7. Demographics
8. BREAK for second part
9. Return information treatment (review)
10. Draw of risky asset return, realized satisfaction
11. Investment decision II
12. Financial literacy
13. Draw of return II and realized satisfaction II

C.II. Transcript of instructions and questions

1. Comprehension

1.1. Example 1 [exp. 2,3 only]

Q) You invest $1, 000 in the risky asset. The risky asset makes a return of 15%.
How much money do you have in total?

- $1300
- $1150
- $950
- $1050

Q) How much will you earn as a bonus payoff?

- $1
- $1.15
- $10
- $11.50

1.2. Example 1 solved

We want to make sure that all you understand the payoff mechanism:
You invest $1,000 in the risky asset. The risky asset makes 15%.

I) How much money do you have in total?
Your answer was correct / Your answer (XXX) was wrong. The correct answer is $1150.
II) How much will you earn as a bonus payoff?
Your answer was correct / Your answer (XXX) was wrong. The correct answer is $11.50.

1.3. Questions [exp 2 only]

Q) If you make $1,200 in this study and will be selected for additional payoff, how much real money will you receive as a bonus payoff (additionally to the $0.51 for participating)?

- $0.51
- $0.60
- $1.20
- $12.00

Q) If you invest more in the risky asset, your payoff variability will be

- Higher
- Lower
- The same

1.4. Example 2 [exp. 2 only]

You invest $500 in the risky asset. The return of the risky asset is -10%. You still have the $500 you have not invested. Q) How much money do you have in total?

- $1000
- $900
- $950
- $1050

Q) How much will you earn as a bonus payoff?

- $10
- $9
- $9.50
- $0.90

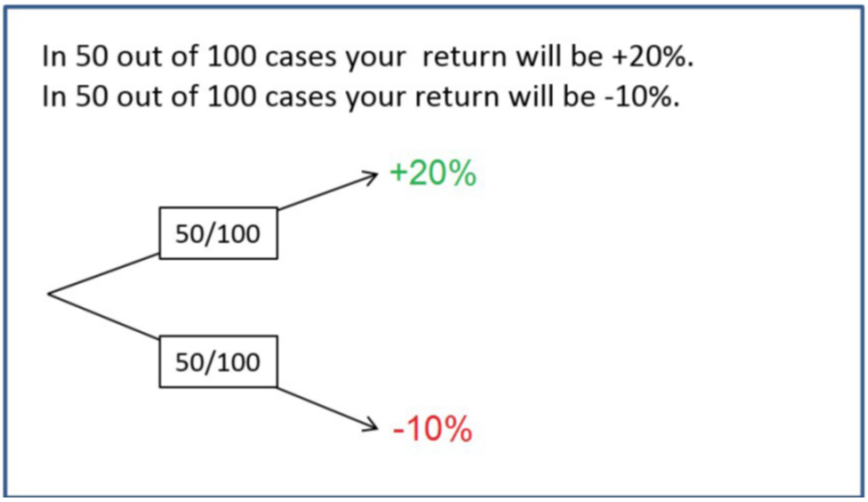
1.5. Example 2 solved

Again, we want to make sure that you understand the investment mechanism:
You invest $500 in the risky asset. The return of the risky asset is -10%. You still have the $500 you have not invested.
I) How much money do you have in total?
Your answer was correct / Your answer (XXX) was wrong. The correct answer is $950.
II) How much will you earn as a bonus payoff?
Your answer was correct / Your answer (XXX) was wrong. The correct answer is $9.50

2. Return Information Treatments

2.1. Descriptive-Simple [experiment 1 only]

The asset is risky and can have different returns.
The following graph shows you the potential returns of the asset and their corresponding probabilities.

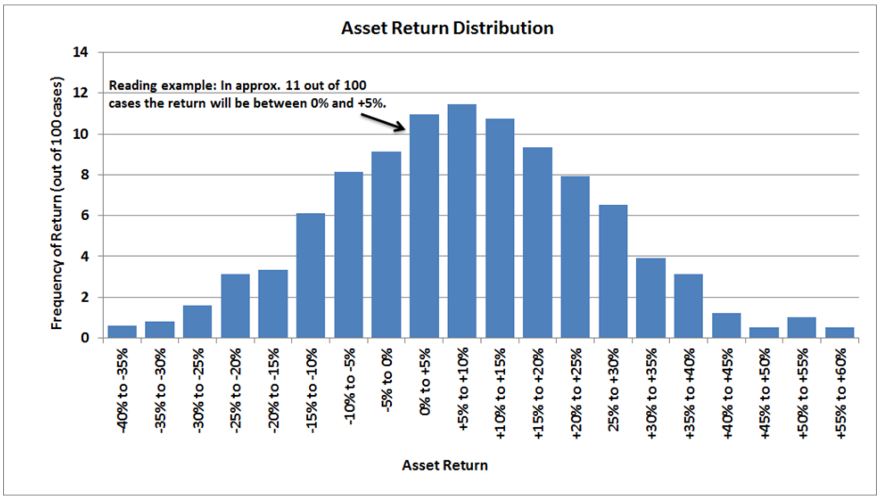


2.2. Descriptive-Complex [experiment 1]

The asset is risky and can have different returns.

The following graph shows you potential returns of the asset. On the x-axis you see return categories and on the y-axis you see the frequency with which these returns occur.

The higher the respective bar, the more likely such a return will be. Please look at the returns carefully as you will have to decide how much to invest in this asset.

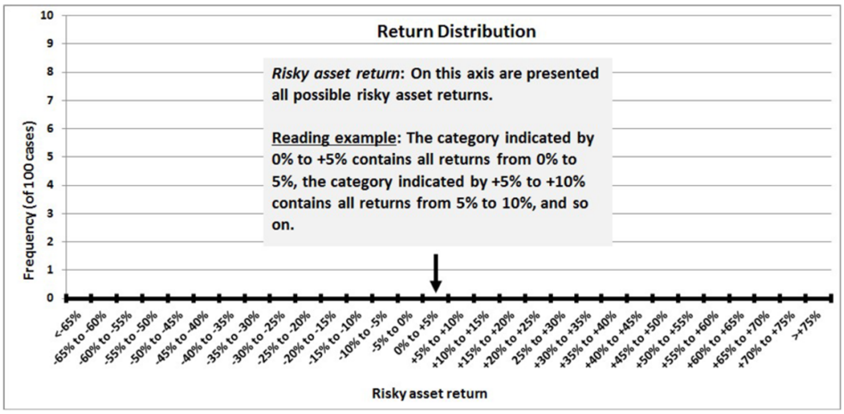


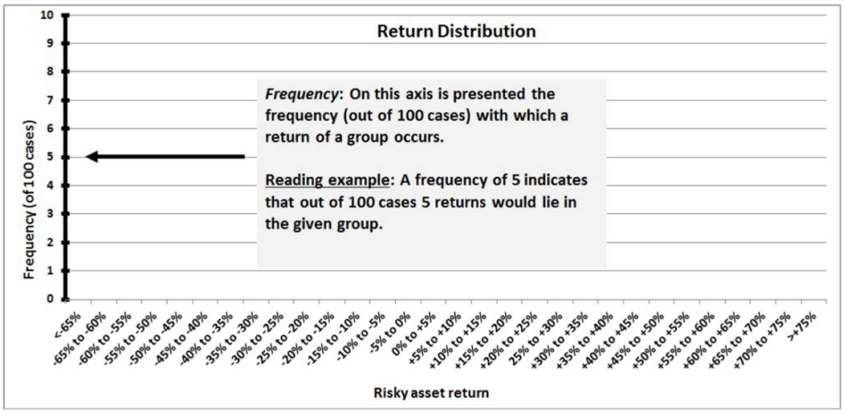
2.3. Descriptive-Complex [experiment 2]

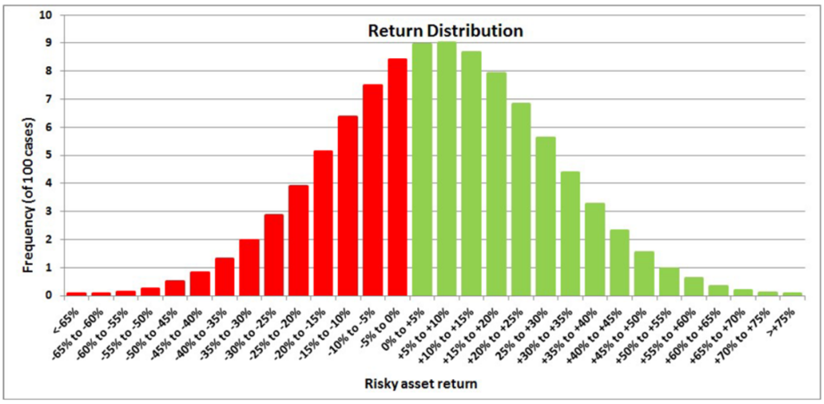
The asset is risky and can have different returns.

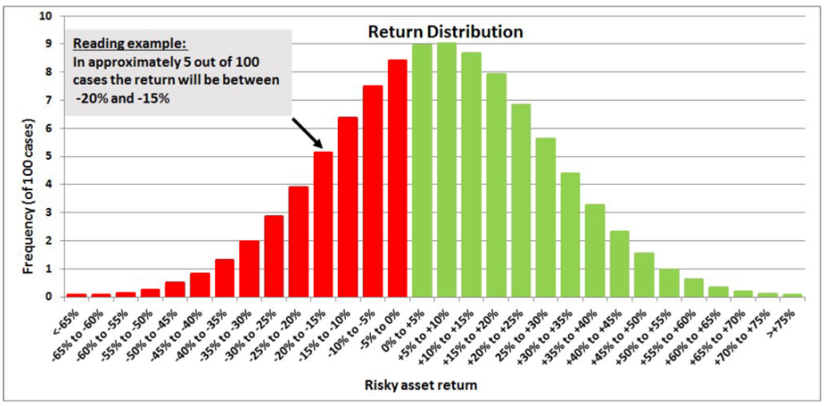
The following graph shows you potential returns of the asset. On the x-axis you see return categories and on the y-axis you see the frequency with which these returns occur.

The higher the respective bar, the more likely such a return will be. Please look at the returns carefully as you will have to decide how much to invest in this asset.


Return Distribution
Risky asset return: On this axis are presented all possible risky asset returns.
Reading example: The category indicated by 0% to +5% contains all returns from 0% to 5%, the category indicated by +5% to +10% contains all returns from 5% to 10%, and so on.
Frequency (of 100 cases)
Risky asset return

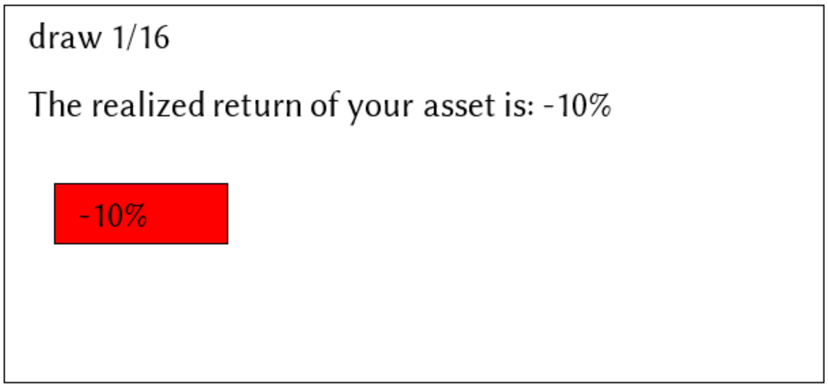

Return Distribution
Frequency: On this axis is presented the frequency (out of 100 cases) with which a return of a group occurs.
Reading example: A frequency of 5 indicates that out of 100 cases 5 returns would lie in the given group.
Frequency (of 100 cases)
Risky asset return


Return Distribution
Frequency (of 100 cases)
Risky asset return


Return Distribution
Reading example:
In approximately 5 out of 100 cases the return will be between -20% and -15%
Frequency (of 100 cases)
Risky asset return

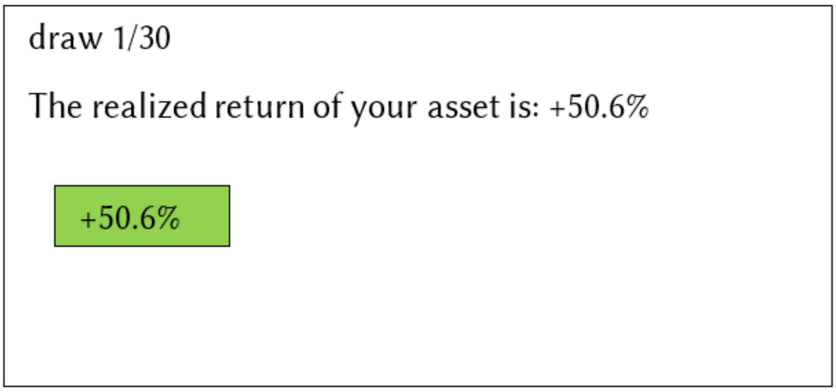
2.4. Experience-Simple [experiment 1 only]

The asset is risky and can have different returns. You will now see 16 potential returns of the asset after each other. These 16 returns are randomly drawn out of the asset's return distribution, and they are representative for the asset's return distribution. The more often a return occurs, the more likely it is.
Please look at the returns carefully as you will have to decide how much to invest in this asset.
Click forward to draw the 16 random returns...



2.5. Experience-Complex [exp. 1 and 2]

The asset is risky and can have different returns. You will now see 30 potential returns of the asset after each other. These 30 returns are randomly drawn out of the asset's return distribution, and they are representative for the asset's return distribution. The more often a return occurs, the more likely it is.
Please look at the returns carefully as you will have to decide how much to invest in this asset.
Click forward to draw the 30 random returns...



2.6. Risky asset die [experiment 3]

Possible retuns of the investment asset
The asset is risky and can have six different returns. All of these six return are equally likely to

occur.

The following table shows you potential returns of the asset. In the first column you see the 6 possible returns. The next column shows their respective probability. The last column illustrates the corresponding outcome of throwing a single die. Please look at the returns carefully as you will have to decide how much to invest in this asset.

Return	Probability	Die Outcome
-20%	1/6	⚀
-10%	1/6	⚁
0%	1/6	⚂
10%	1/6	⚃
20%	1/6	⚄
30%	1/6	⚅

3. Riskiness and Expectations

What annual return of the risky asset do you expect on average? (exp. 3 only)

Q) Please provide your best guess in %. (Write your guess without the % behind it, and without decimals.)

[num input]

Q) In how many out of 100 cases would you expect the return to be negative (lower than 0%)?

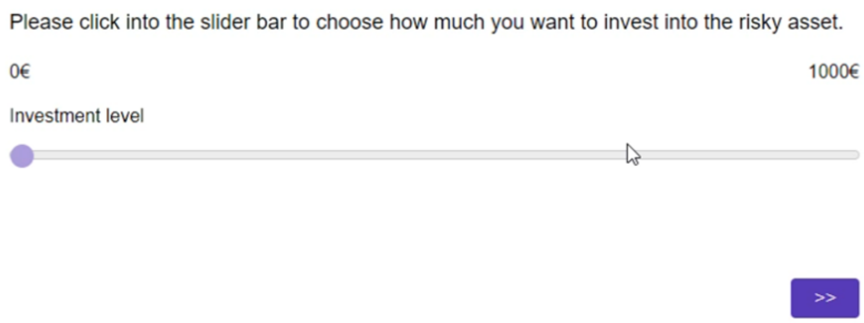
[num input]

Q) How risky do you perceive this investment to be?

[9-point Likert scale from “not risky at all” to “very risk” with “neutral” at radio button 5]

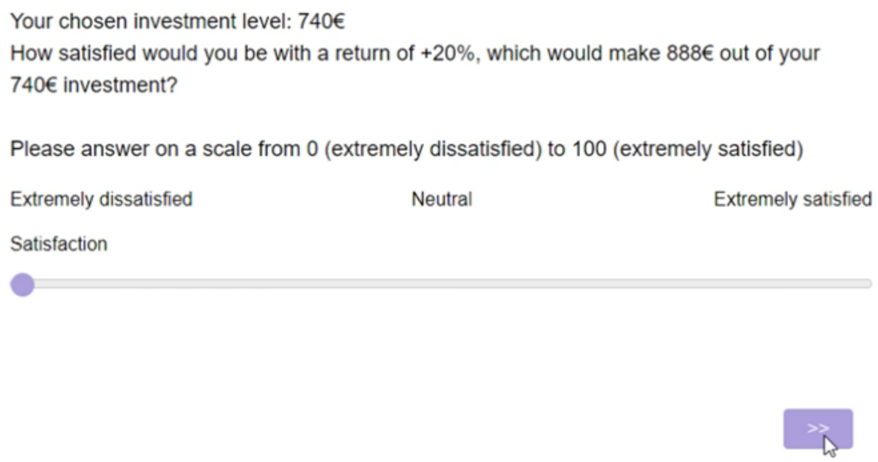
4. Investment decision

Allocation [1000 point continuous slider from $0 to $1000]



5. Anticipated Satisfaction [exp. 1,2]

Your chosen investment level: $X Q) How satisfied would you be with a return of +/-20%, which would make $X out of your $X investment? Please answer on a scale from 0 (extremely dissatisfied) to 100 (extremely satisfied) [100 point continuous slider]



6. Anticipated Satisfaction [exp. 3]

Q) How satisfied would you be with a return of +/-20%? Please answer on a scale from 0 (extremely dissatisfied) to 100 (extremely satisfied) [100 point continuous slider]

7. Financial Literacy

Thank you for your allocation decision. To complete the study, we will ask you some additional questions.

Q) Imagine that the interest rate on your savings account was 1% per year and inflation was 2% per year. After 1 year, would you be able to buy:

- More than today with the money in this account
- Exactly the same as today with the money in this account
- Less than today with the money in this account
- Don't know
- Refuse to answer

Q) Do you think that the following statement is true or false? "Bonds are normally riskier than stocks."

- True
- False
- Refuse to answer
- Don't know
- Refuse to answer

Q) Considering a long time period (for example, 10 or 20 years), which asset described below normally gives the highest return? [exp. 1,2 only]

- Savings accounts
- Stocks

- Bonds
- Don't know
- Refuse to answer

Q) Normally, which asset described below displays the highest fluctuations overtime? [experiment 1 only]

- Savings accounts
- Stocks
- Bonds
- Don't know
- Refuse to answer

Q) When an investor spreads his money among different assets, does the risk of losing a lot of money:

- Increase
- Decrease
- Stay the same
- Don't know
- Refuse to answer

Q) Do you think that the following statement is true or false? "If you were to invest $1,000 in a stock mutual fund, it would be possible to have less than $1,000 when you withdraw your money."

- True
- False
- Don't know
- Refuse to answer

Q) Do you think that the following statement is true or false? "A stock mutual fund combines the money of many investors to buy a variety of stocks." [experiment 1 only]

- True
- False
- Don't know
- Refuse to answer

Q) Do you think that the following statement is true or false? "After age 70 1/2, you have to withdraw at least some money from your 401(k) plan or IRA." [experiment 1 only]

- True
- False
- It depends on the type of IRA and/or 401(k) plan
- Don't know

- Don't know
- Refuse to answer

Q) Do you think that the following statement is true or false? "A 15-year mortgage typically requires higher monthly payments than a 30-year mortgage, but the total interest paid over the life of the loan will be less." [experiment 1 only]

- True
- False
- Don't know
- Refuse to answer

Q) Suppose you have $100 in a savings account and the interest rate is 20% per year and you never withdraw money or interest payments. After 5 years, how much would you have in this account in total?

- More than $200
- Exactly $200
- Less than $200
- Don't know
- Refuse to answer

Q) Which of the following statements is correct? [experiment 1 only]

- Once one invests in a mutual fund, one cannot withdraw the money in the first year
- Mutual funds can invest in several assets, for example invest in both stocks and bonds
- Mutual funds pay a guaranteed rate of return which depends on their past performance
- None of the above
- Don't know
- Refuse to answer

Q) Which of the following statements is correct? If somebody buys a bond of firm B: [experiment 1 only]

- He owns a part of firm B
- He has lent money to firm B
- He is liable for firm B's debts
- None of the above
- Don't know
- Refuse to answer

Q) Suppose you owe $3,000 on your credit card. You pay a minimum payment of $30 each month. At an annual percentage rate of 12% (or 1% per month), how many years would it take to eliminate your credit card debt if you made no additional new charges? [experiment 1 only]

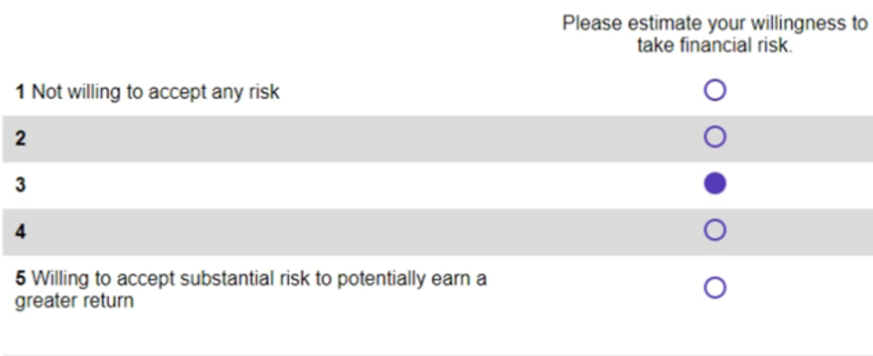
- Less than 5 years
- Between 5 and 10 years
- Between 10 and 15 years
- Never
- Don't know
- Refuse to answer

Q) Out of 1,000 people in a small town 500 are members of a choir. Out of these 500 members in a choir 100 are men. Out of the 500 inhabitants that are not in a choir 300 are men. What is the probability that a randomly drawn man is a member of the choir? [experiment 1 only]
Please indicate the probability in percent. This means that you should not use any commas or dots. [num Input]
Q) Imagine we are throwing a five-sided die 50 times. On average, out of these 50 throws how many times would this five-sided die show an odd number (1, 3 or 5)? [experiment 1 only] [num Input]
Q) Imagine we are throwing a loaded die (6 sides). The probability that the die shows a 6 is twice as high as the probability of each of the other numbers. On average, out of these 70 throws how many times would the die show the number 6? [experiment 1 only] [num Input]
Q) In a forest 20% of mushrooms are red, 50% brown and 30% white. A red mushroom is poisonous with a probability of 20%. A mushroom that is not red is poisonous with a probability of 5%. What is the probability that a poisonous mushroom in the forest is red? [experiment 1 only] [num Input]

8. Demographics

8.1. Risk tolerance

[5-point Likert scale from "not willing to accept any risk" to "willing to accept substantial risk to potentially earn a greater return"]



Q) Do you own stocks or an equity mutual fund?
[Yes,No]
Q) If you are not invested in the stock market (via stocks or mutual funds). Why? [experiment 3 only]

- It is too risky.
- I don't have any savings.
- I don't trust the stock market.
- I did not have the time to select suitable investment products.
- other

Q) How interested are you in the stock market or in financial markets in general? [experiment 3 only]
[5-point Likert scale from "not interested at all" to "very interested"]
Q) Are you generally interested in stock or financial markets? [experiment 1,2 only]
[Yes,No]
Q) Are you male or female?
[male, female]

9. Numerical literacy [experiment 2 only]

Q) Two iron plates weigh 55lbs in total. The smaller plate weighs 50lbs less than the heavier one. How many lbs does the small plate weigh?
[num input]
Q) 3 workers need 3 days to produce 3 guitars. How many days would it take 10 workers for 10 guitars?
[num input]
Q) New computers get faster and faster. Every year, their speed doubles. If it takes 10 years from now for computers to reach a particular speed, how long will it take for computers to be half as fast as this particular speed?
[num input]
Q) Peter is running in a race and passes the person in 2nd place, what place would Peter be in?
[num input]

10. Draw of Return and Experienced Satisfaction

We will now conduct a market simulation in order to draw the return of your investment based on the underlying distribution.
Click next to see the outcome of your investment into the risky asset...
Your investments' return is
XX%
Q) How satisfied are you with the outcome?
Please answer on a scale from 0 (extremely dissatisfied) to 100 (extremely satisfied) [100 point continuous slider] (see slider for anticipated satisfation).

11. Investment decision II [experiments 1,2]

Q) If you could now invest into the same asset for a second time - how would you choose?
Please click into the slider bar to choose how much you want to invest into the risky asset

[1000 point continuous slider from 0€ to 1000€]

12. Investment decision II [experiments 3]

Now follows a second, independent, investment period. Make your decision for this new investment period:

You are again endowed with 1000€. The outcomes of the first and second year are independent, so the risky asset's outcome for the following period does not depend on the old return. The return for the second year will be drawn from the same return distribution, that you saw again at the beginning of this part.

At the end of this study, we will randomly pick one of the two years by a coin toss, which will determine your payout (so either year 1 or year 2).

Q) How much do you invest into the same risky asset for the coming second year?

Allocation [1000 point continuous slider from 0€to 1000€]

13. Draw of Return II and Experienced Satisfaction II [exp. 3 only]

The return of your second investment is

XX%

Q) How satisfied are you with the outcome?

Please answer on a scale from 0 (extremely dissatisfied) to 100 (extremely satisfied)

[100 point continuous slider]

Chapter II.

Are Financial Advisors (Good) Match-Makers?

Are Financial Advisors (Good) Match-Makers?

Matthias Rumpf[*] and **Thomas Otter**[†]

July 7, 2020

Abstract

In this paper we examine the ability and heterogeneity of lay and professional advisors in matching investor demographics, such as age and income, into risky asset portfolio shares. To this end, we use portfolio recommendations collected in a web-based experiment from among 424 independent professional financial advisors, along with 450 regular people "from the street", i.e., lay advisors. Clients' risk tolerance, age, and income appear to be the most important factors in the advisors' portfolio rules, with clients' wealth only playing a secondary role. Moreover, based on the estimated heterogeneity, advisors seem to disagree primarily on how to map financial wealth into risky asset shares. Very young, risk tolerant, or rich clients seem to receive congruent advice to hold large risky asset shares, while poor, highly risk averse and old clients are more likely to experience greater heterogeneity in the recommendations made to them. Our estimates suggest that advisors' age, long-term return expectations, and risk tolerance directly push up the recommended risky share. Comparing professional and lay advisors, we find that professionals add value by incorporating client characteristics more strongly and reliably into their portfolio recommendations. Also, there is no evidence that professionals integrate their personal beliefs or own investment strategies any more than lay advisors.

*. Goethe University Frankfurt
†. Goethe University Frankfurt

Introduction

Three out of four Germans are able to save at least part of their income; 45% save to their current account, while 21% hoard cash at home. By shunning investment opportunities savers forego between €20 billion and €30 billion in returns per year, according to a survey sponsored by Postbank and Kantar Emnid (2017). If they do invest, the majority of Germans rely on financial advisors; 67% of participants in another survey of Union Investment and Forsa (2018) stated that they rely on their advisors to find alternative investment opportunities in the current low interest rate environment. A stunning 40% of participants revealed that they have insufficient knowledge in matters of investment and 48% are not aware that stock investments yield the highest long-term expected returns. It is therefore not surprising that only 12% of Germans between the age of 15 and 59 invest in the stock market, either directly or via funds (Deutsches Aktieninstitut (2017)).

Evidently, savers are urgently in need of proper financial advice. Most of them still need to be convinced of the benefits of stock market participation, and a smaller group of investors requires help in finding suitable investment opportunities in difficult market situations. For both groups, it is paramount to know the value added by financial advisors. The aim, in the case of the first group, is to eventually overcome concerns about consulting a financial advisor, while the second group needs to evaluate whether their advisors provide additional knowledge and insights.

In this study, we compare portfolio recommendations of professional advisors and regular people to evaluate the benefits of financial advice. While financial advice has been found to improve diversification in client portfolios (see, for example, Bluethgen et al. (2008), Hackethal et al. (2012)), we might expect mis-aligned incentives to encourage advisors to sell more risky and costly high-commission products (Inderst and Ottaviani (2009)). And, indeed, in a large audit study, advisors did not mitigate obvious behavioral biases of their clients, but catered to detrimental mistakes such as return chasing and tilted even passive low cost portfolios towards inefficient, actively managed high-cost products (Mullainathan et al. (2012)). Nevertheless, recent research indicates that financial advisors are not necessarily intentionally abusing their clients' trust, but might be steered by 'misguided beliefs'. Using a large brokerage data set featuring more than 10,000 financial advisors and 800,000 clients, Linnainmaa et al. (2018) find that clients mirror their advisors' investment mistakes who deteriorate their own portfolios' returns by trading frequently, chasing returns, under-diversifying and investing in costly actively managed funds. In an earlier study on the same data, Foerster et al. (2017) report that receiving advice from a bottom-decile compared to a top-decile advisor results in 1.2% lower gross returns. They further state that advisor-fixed effects, which reflect differences in advisor performance, are closely linked to the advisors' age and risk tolerance.

An important aspect that has thus far received less attention is the matching of client characteristics with optimal portfolio allocations. Filling this gap, we add to the large body of literature on the socio-demographic determinants of portfolio choice.[1] Since Merton (1969), Mossin (1968), and Samuelson (1969) (MMS) pioneered the theory of modern life-cycle portfolio choice theory and the normative prescription that investors should participate in stock markets at all ages, portfolio models have gradually increased in complexity.[2] All models and their normative theory share a common challenge in that they are notoriously hard to test on real portfolio data. Take as an example the wealth elasticity of the risky asset portfolio share. Early papers, such as by Friend and Blume (1975), have found evidence of constant relative risk aversion in the context of portfolio choice.[3] Contrasting with previous works, Calvet and Sodini (2014) find evidence of decreasing relative risk aversion; in other words, increases in financial wealth are associated with higher risk-taking. They provide an unprecedented level of control using data on Swedish twins.

In our paper, we test the effects of financial wealth and other demographic variables in an experimental portfolio allocation task that overcomes previous empirical challenges by ensuring that all household characteristics are exogenous. Recovering financial advisors allocation rules from data on client portfolios is just as challenging as testing normative theory. For example Foerster et al. (2017) showed that clients' demographic variables, in particular risk tolerance and age, explain 12% of variation in advised brokerage accounts. Yet, it remains impossible to distinguish whether the underlying portfolio rules are prescribed and enforced by the advisors or the clients themselves.

Entering a clinic as a patient to receive medical attention, we certainly hope that our doctor takes at least some of our personal characteristics into account when prescribing a treatment. Likewise, as investors, we hope that financial advice goes beyond, say, recommending a replication of the

1. See, for example, Mankiw and Zeldes (1991) and Haliassos and Bertaut (1995) on stock market participation; Bodie and Dwight B. Crane (1997), Heaton and Lucas (2000a), and Bertaut and Starr-McCluer (2000) on portfolio allocation.

2. In a seminal paper, Cocco et al. (2005) estimate one of the most comprehensive portfolio choice models found in the literature. In comparison with the benchmark models of MMS their realistically calibrated life-cycle model includes risky, non-tradable labor income, borrowing constraints, mortality risk, a penalty for default, uncertain retirement income, a bequest motive, and recursive preferences that allow to decouple time preferences and risk aversion. It predicts a moderately declining risky asset portfolio share over the life-time, induced mainly by the labor income process. Human capital, which to some extent, serves the purpose of a safe asset, declines as the investor ages.

3. For instance, Calvet and Sodini (2014) argue that the risky asset share increases with wealth. The controversy arises because the positive correlation observed in cross-sectional data does not necessarily entail an unambiguous causal relation. For instance, the risky asset portfolio share could increase if wealth increases (ceteris paribus), which implies decreasing relative risk aversion. But an alternative explanation is also the positive correlation of socio economic status (e.g. income, education and occupation) and risk tolerance (see Calvet and Sodini (2014)). Wealth is positively correlated with income, education, occupation and possibly other unobserved investor characteristics which are by themselves closely related to the investors' risk preferences. In a cross-sectional estimation of asset allocations it is impossible to determine whether richer households hold riskier portfolios or whether households are richer because they are less risk averse (Guiso and Sodini (2013)). Solutions were proposed by Chiappori and Paiella (2011) who exploit panel data by considering time variations in portfolios, wealth, and other characteristics. And Brunnermeier and Nagel (2008), who use instrumental variables on financial wealth like inheritance or income growth. Unfortunately these solutions are not perfect, first, panel data would need to be very frequent and accurate so that the researcher could distinguish changes in the portfolio allocation due to passive variation, by market movements, from active changes in allocation. Furthermore active changes in portfolio allocation might be subject to infrequent rebalancing, which is referred to as portfolio inertia.

advisor's personal portfolio. In fact almost everyone[4] expects financial advisors to tailor recommendation to their individual situation and personal needs. Using an experimental approach and a sophisticated estimation routine, we go beyond the analyses of Foerster et al. (2017) who find that advisor fixed effects are driven by the advisors own risk allocations. We are not only able to identify drivers of risk taking such as the clients age and income, but to explore if the characteristics of advisors can be associated with heterogeneity in the mapping of client characteristics into optimal portfolio recommendations. In particular, we analyze differences in the portfolio choice between regular investors and professional advisors.

In this paper, we use data on portfolio recommendations collected in a web-based experiment among 424 independent financial advisors that we refer to as "professional advisors" and 450 regular people recruited "from the street" referred to as "lay advisors". Each participant stated optimal risky asset portfolio shares for five virtual clients described by socio-economic household characteristics. Among the recommendations, we elicit the participants' return expectations, risk preferences, demographic variables and information about their own portfolios. The set-up of our experiment overcomes the difficulties encountered in traditional datasets. As discussed, endogeneity poses the greatest impediment when estimating determinants of the risky asset share – a problem we solve by making household characteristics completely exogenous. Furthermore, observational data makes it impossible to disentangle the decisions and thus the impact of clients and advisors on the portfolios under review. In our experiment, hypothetical clients introduced to participating advisors show, for the most part, no or very little correlation across characteristics. We include a variety of advisor characteristics and introduce advisor-specific random intercepts. Since we observe five recommendations for each participant, we are be able to control for unobserved heterogeneity. It is therefore plausible to assume that our estimates are free from potential endogeneity and any selection bias, for example, a systematic matching of clients with advisors.

We aim to reveal how households believe they should best allocate retirement savings. If households have a clear idea of optimal portfolio allocation that financial advisors we observe could not improve upon, the financial industry should focus on alleviating obstacles that keep households from participating more intensively and optimally in financial markets. This could be achieved by providing standardized, comprehensive and low-cost products instead of selling time-intensive personal advice.

Exploring the determinants of portfolio recommendations, we use a Markov chain Monte Carlo (MCMC) sampler to estimate the coefficients of a parsimonious Bayesian linear mixed model that regresses optimal risky asset share recommendations on advisor and household profile characteristics. It is mixed in the sense that it includes both random and fixed effects[5], and parsimonious by implementation of multiple variable selection steps that adaptively reduce the effective parameter space in our model. We extend the model of Frühwirth-Schnatter and Tüchler (2008), who introduced selection on the random effects covariance matrix in Bayesian linear mixed models.

4. 97% of participants in a survey among 1,026 German adults (Net Fonds and Toluna (2015))

5. Note that we use the term fixed effects to refer to the set of non-random effects, i.e. non-random regression coefficients, in our estimation model as in, e.g., Frühwirth-Schnatter and Tüchler (2008).

We add a second selection step on the large number of potential fixed effects in our model. The resulting model allows us to predict risky asset share recommendations for an array of household profiles for each participating advisors, enabling us to evaluate heterogeneity in advice and compare recommendations of professional and lay advisors.

Our contribution is three-fold. First, we identify the determinants of the participants' portfolio allocation rules and test normative portfolio predictions. Our estimates show that clients' risk tolerance, age, and income are the most important factors in the advisors' portfolio rules. Surprisingly, judging by the size of the effect, wealth – whether financial or real estate – plays only a secondary role. Second, we add to the literature on financial advice by providing an in-depth analysis on the heterogeneity of portfolio recommendations. We find that, in addition to the advisors' own risky allocation, as emphasized by Linnainmaa et al. (2018), advisors' age, long-term return expectations, and risk tolerance directly increase the recommended risky share by roughly 6% to 9% for an average client following an increase of close to one standard deviation in the corresponding advisor characteristic. When incorporating the information on household demographics, advisors seem to disagree primarily on how to map financial wealth into risky asset shares. Advisors' age, return expectations, and risk tolerance identify advisor groups with heterogeneous opinions of how client characteristics should be mapped to achieve optimal allocation. Concerning differences in recommendation heterogeneity across households, we find that very young, risk tolerant or rich clients appear to receive congruent advice to hold large shares of risky assets, while poor, very risk-averse and old clients are more likely to face larger heterogeneity in recommendations. Comparing professional and lay advisors, we find that professionals add value by incorporating client characteristics more strongly and reliably into their portfolio recommendations. Testing for differences in "one size fit all" heuristics (a term coined by Foerster et al. (2017)) between professionals and lay advisors, we find no evidence of professionals incorporating their beliefs or own investment strategies more strongly into recommendations than lay advisors. Third, and finally, we demonstrate the application of a parsimonious Bayesian variable selection method that allows us to investigate heterogeneous effects on a large set of interactions using a large N small T dataset. One approach that is certainly applicable to a range of different topics, e.g. the analysis of heterogeneity in treatments of different patients across medical practitioners.

In section 1, we outline the experiment's design and the resulting data. We derive an appropriate portfolio allocation rule, and introduce our model and estimation approach in section 2. Section 3 summarizes normative theory on portfolio choice. Results are presented in section 4, in which we examine the estimation results to identify important determinants of advisors' allocation rules, the sources of heterogeneity, and the differences between professional and lay advisors. Section 5 concludes.

1. Experimental design and data

In a web-based survey, we collected data on the risky asset share portfolio allocation recommendations on long-term retirement investments. We recruited participants from two groups – independent financial advisors (professional advisors) and regular "people from the street" (lay advisors). In April 2015, we approached 10,000 independent, non-bank advisors directly by mail. Their postal addresses were obtained from the register of finance and insurance brokers, which is publicly accessible and provided by the German chamber of commerce.[6] The survey for professionals ran from 13 April to 27 April 2015. Around 800 visits resulted in 424 full responses. The complementary survey among non-professionals, recruited by a market research service, ran from 15 April to 5 Mai 2015 and produced 450 complete responses. Professional advisors were compensated by being given access to the results of a set of questions regarding the industry of independent financial advice. Non-professionals were compensated by the market research service provider.

Table 1.
Experiment set-up

Section #	Question section title
1.	Screening / Financial literacy (only lay advisors)
2.	Elicitation of return expectations
3.	Elicitation of risk preferences
4.	Portfolio experiment: Allocations for household profiles I-V
5.	Elicitation of participant demographics

The general setup of the portfolio experiment is illustrated in Table 1. The lay advisor group was screened for age and a financial literacy questionnaire to include only participants capable of expressing basic opinions and beliefs on investment decisions. Lay advisors had to be at least 25 years old, answer questions 2.1-2.3 correctly, and question 2.4 with yes, see Appendix B for details on the survey. Out of 6507 logins to the survey, 696 passed the admission and completed the survey, 450 participants completed the questionnaire with valid entries. We excluded, for example, obviously erroneous entries that included only "0" in all numeric answers. The rest of the experiment was almost identical for both groups. First, we elicited return expectations and risk preferences, then participants were asked to give an optimal portfolio allocation recommendation for five virtual clients. After seeing the last profile, participants stated their own optimal and actual allocation as well as information on their own demographic and financial situation. A detailed description of the experiment together with an English translation of the original German version is included in Appendix B.

6. https://www.vermittlerregister.info/

Anlegerprofil	
Persönliche Merkmale	
☐ Alter	59
☐ Geschlecht	männlich
☐ Verheiratet	ja
☐ Kinder	0
☐ Bildung	Hochschulstudium
☐ Beruflicher Status	angestellt
☐ Risikoeinstellung	2
Finanzielle Merkmale	
☐ Einkommen	15.000 EUR
☐ Davon sicheres Einkommen	60%
☐ Ausgaben für Lebensmittel pro Jahr	5.500 EUR
☐ Immobilienvermögen	125.000 EUR
☐ Hypothekenkredit	5.000 EUR
Merkmale zur Kapitalanlage	
☐ Erfahrung mit Aktieninvestments	1 bis 3 Jahre
☐ Anlagehorizont	Geldanlage für den Ruhestand
☐ Anlagebetrag	15.000 EUR

Figure 1.
Household profile display.
The screenshot above shows a household profile as presented to participants during the experiment. Each characteristic provides a detailed description when hovering over it with the mouse. Respondents were asked to check the boxes on the left to indicate which characteristics they considered applicable.

Virtual client profiles

The core of our experiment is the portfolio allocation task. Before the allocation task, a comprehensive description of each demographic household characteristic was provided. Each description could be reviewed by the participant when hovering over it with the mouse pointer during the allocation task. Each participant was provided with five hypothetical investor profiles. Each hypothetical investor was presented to participants in the form of a table with information on 14 characteristics (see Figure 1 for an example). To track the characteristics that enter the participants' decision rule, they were asked to check boxes next to the characteristics for at least the first profile. Around 80% of participants checked at least one box. Finally, participants had to state the amount in euro that they felt represented optimal allocation to risky assets, while the percentage portfolio weights were calculated and displayed next to the entry box in real time.

The virtual client profiles were combined efficiently to maximize information for a linear regression model. This was achieved by utilizing the R-package "algdesign". With 14 variables it is not possible to generate a full factorial design matrix even with very parsimonious discretization. Therefore, we generated two million randomly drawn profiles from which the algorithm selected sets of five profiles for each participant. Income, financial wealth and real estate wealth were randomly generated to match a distribution similar to the data in the German Socio Economic Panel (GSOEP) for households with wealth exceeding €10,000. Summary statistics on the profiles shown to professionals and lay advisors are presented in Table 2. Comparing means, and quantiles which are not shown in the table, both groups have faced a similar distribution of client profiles. Only the investment amount of professional advisors seem to show higher values due the random sampling. See Appendix B for details on the quasi-random generation of household profiles and the description of characteristics provided to participants.

Table 2.
Household profiles - summary statistics on demographic variables
The table collects summary statistics on household profiles shown to participants during the portfolio allocation task. The first two sets of columns report statistics separately for the advisor groups. The last two columns show the overall mean and standard deviation. Statistics in rows one to five provide information on the investment amount down to food spending and is shown in thousand euro. Safe income is a fraction (of total income). All other variables except for the number of kids, risk tolerance, and age are binary indicators. For more information, see Table A.3 in Appendix A and the survey set-up in Appendix B. which provide further details on the household characteristics.

	Profess. Advisor				Lay Advisor				Overall	
	N	Mean	Min	Max	N	Mean	Min	Max	Mean	SD
Inv.Amount	2105	113.7	10	1000	2250	106.4	10	1000	109.9	199.6
Income	2105	43.8	5	150	2250	42.7	5	150	43.3	32.1
RE.Wlth	2105	217.0	0	1500	2250	209.7	0	1500	213.2	283.2
RE.Dbt	2105	17.0	0	400	2250	15.9	0	350	16.4	39.9
Food Expnd.	2105	13.1	1	72	2250	12.9	1	71	13.0	11.2
Safe Inc.	2105	85.4	50	100	2250	86.0	50	100	85.7	18.6
RiskTol.Prov.	2105	0.8	0	1	2250	0.8	0	1	0.8	0.4
RiskTol.	2105	2.4	0	5	2250	2.4	0	5	2.4	1.8
Inv.Exp.<1yr	2105	0.3	0	1	2250	0.3	0	1	0.3	0.5
Inv.Exp.>3yrs	2105	0.3	0	1	2250	0.3	0	1	0.3	0.5
Age	2105	45.5	20	75	2250	45.4	20	75	45.5	15.8
Male	2105	0.5	0	1	2250	0.5	0	1	0.5	0.5
Married	2105	0.7	0	1	2250	0.7	0	1	0.7	0.4
Kids	2105	1.5	0	4	2250	1.6	0	4	1.6	1.5
Prof.Train.	2105	0.2	0	1	2250	0.3	0	1	0.2	0.4
ALevels	2105	0.3	0	1	2250	0.3	0	1	0.3	0.4
College	2105	0.3	0	1	2250	0.3	0	1	0.3	0.4
Retired	2105	0.1	0	1	2250	0.1	0	1	0.1	0.3
SelfEmpl.	2105	0.1	0	1	2250	0.1	0	1	0.1	0.3

Participants

We refer to the two groups of participants as "professional advisors" (independent advisors that received an invitation via mail) and "lay advisor" ("people from the street", recruited by a market research firm). The latter sample was screened against individuals that are not interested in investments or fail to pass a basic financial literacy test. We compare the participant groups in Table 3. We see that around 90% of professional financial advisors are male, whereas almost 40% of lay advisors are female. Professionals advisors are richer and more risk tolerant, but hold similar return expectations with respect to stock investments. The most salient difference is in the share of risky assets. Professionals hold 52% while the share of risky assets in the lay advisors' portfolio represents an average of only 32%. Details on the corresponding survey questions are provided in Appendix B.

Table 3.
Participant demographics

The table reports summary statistics on participating advisors. The first two sets of columns report statistics separately for both advisor groups. The last two columns show the overall mean and standard deviation. Male and college are binary indicators, income and wealth are in thousand euro, and current allocation is the advisors' own private risky asset share in percent. Return expectations are beliefs on an average annual return above 10% in percentage probabilities. The risk tolerance is measured via the "bomb-game" on a scale from 1 to 99 and patience from 1 to 10. For more information, see Table A.2 in Appendix A and the survey setup in Appendix B, which provide more details on the data collected on advisors.

	Profess. Advisor				Lay Advisor				Overall	
	N	Mean	Min	Max	N	Mean	Min	Max	Mean	SD
Age	421	46.3	23	72	450	47.0	25	76	46.7	11.9
Male	421	0.9	0	1	450	0.6	0	1	0.8	0.4
College	421	0.4	0	1	450	0.5	0	1	0.5	0.5
Income	421	70.0	13	250	450	43.5	5	250	56.3	39.8
Real Estate Wealth	421	310.6	0	3000	450	117.5	0	3000	210.8	463.5
Financial Wealth	421	124.8	5	1000	450	61.4	5	1000	92.0	163.6
Own Allocation	421	52.7	0	100	450	30.2	0	100	41.1	32.1
Return Expectations 10yrs	421	23.4	0	95	450	20.9	0	99	22.1	20.5
Return Expectations 1yr	421	22.5	0	90	450	22.9	0	99	22.7	20.0
Risk Tol. Bomb Game	421	41.6	1	99	450	34.6	1	99	38.0	25.8
Patience	421	6.3	1	10	450	6.2	1	10	6.2	2.1

2. Model development

In this section, we develop our model. First, we derive the portfolio allocation rule from basic theory in section 2.1. We then discuss parsimonious Bayesian inference for the resulting linear mixed model in section 2.2, and present the basics of the corresponding MCMC in section 2.3. Further technical details are provided in Appendix A.

2.1. The portfolio allocation rule

To specify a structural estimation equation, we start from a simple asset allocation rule used by King and Leape (1998).[7] It defines the optimal risky asset portfolio share α_h of household h given its net worth W_h, (absolute) risk aversion A_h, and the risk σ^2 adjusted return of the risky asset R in excess of the riskless return R_f.

$$\alpha_h = \frac{1}{A_h W_H} \frac{R - R_f}{\sigma^2} \tag{1}$$

Substituting the relative risk aversion $\gamma_h = A_h \cdot W_h$ gives:

$$\alpha_h = \frac{1}{\gamma_h} \frac{R - R_f}{\sigma^2} \tag{2}$$

Taking logs, the above equation can be expressed as:

$$\ln \alpha_h = \beta_0 - \ln \gamma_h \tag{3}$$

with

$$\beta_0 = \frac{R - R_f}{\sigma^2}, \quad \gamma_h = A_h W_H$$

If the risk-adjusted risky asset excess return is known, it can be represented by a constant β_0. As a result, the risky asset share only depends on the constant and a measure of household specific risk aversion, γ_h. King and Leape (1998) propose a linear approximation for $\ln \gamma_h$. They assume the household's risk-taking potential depends on observable household characteristics $\boldsymbol{x}_h$ of length d_h that include for example a household's net worth W_h among other variables.[8] Accordingly, we

7. King and Leape (1998) derive the given demand equation for a risky asset and household h, extending a conventional portfolio choice model to capture observed differences in portfolio compositions. They explicitly incorporate the impact of differences in tax rates amongst assets and households and aim at estimating the joint discrete continuous choice of asset holding and portfolio fraction in a multi-assets environment. They provide a very good example concerning the estimation of the determinants of asset demand. We will follow their approach, abstracting from taxes in our general example, altering notation to avoid confusion, and focusing on a two asset case with a risky and a riskless asset.

8. There is substantial evidence that individual risk aversion depends not only on the financial situation (see e.g. Calvet and Sodini (2014)) but also on demographic variables. For example, risk aversion has been found to increase with age (e.g. Guiso and Paiella (2008), Barsky et al. (1997)) and decrease with the level of education (Vissing-Jorgensen (2002), Christiansen et al. (2008)).

replace $\ln \gamma_h$ with the following term (u_h captures unobservable household characteristics).

$$\ln \gamma_h = \boldsymbol{x}_h \boldsymbol{\beta} + u_h \tag{4}$$

Combining (3) and (4) results in a simple estimation equation of the risky asset demand with error term $\epsilon_h = u_h + e_h$, capturing unobserved household characteristics via u_h and the remaining estimation error through e_h

$$\ln \alpha_h = \beta_0 + \boldsymbol{x}_h \boldsymbol{\beta} + \epsilon_h \tag{5}$$

From this point on we use $y = \ln(\alpha)$ to simplify notation:

$$y_h = \beta_0 + \boldsymbol{x}_h \boldsymbol{\beta} + \epsilon_h \tag{6}$$

Modeling portfolio allocation rules of financial advisors

Now we incorporate the role of financial advisors into the household risky asset demand function to derive an optimal portfolio allocation rule. We want to model the advisor's beliefs about the optimal baseline allocation and the risk-taking potential of a client conditional on the characteristics $\boldsymbol{x}_h$. Assuming that different advisors follow different portfolio allocation rules to map household characteristics into a risky asset share, we anticipate two sources of heterogeneity. First, a shift in the constant, i.e. the baseline risky asset allocation, that could be caused for example by differences in market expectations, risk attitudes and other characteristics of the financial advisors, is easily represented by vector $\boldsymbol{x}_a$ of advisor characteristics; each variable in the vector of advisor characteristic is denoted by $x_{a,j}$ with index $j = 1, ..., d_a$. Second, variation on how certain household characteristics should influence the risky asset share across advisors is introduced into our model through observed and unobserved slope heterogeneity:

We capture observed heterogeneity by interacting the entire set of household variables in $\boldsymbol{x}_h$ with each of the advisor characteristics $x_{a,j}$ in $\boldsymbol{x}_a$ resulting in $\boldsymbol{x}_h \times \boldsymbol{x}_a$. We use the the symbol "×" in $\boldsymbol{x}_h \times \boldsymbol{x}_a$, similar to the Cartesian product, to symbolize that the columns in $\boldsymbol{x}_h \times \boldsymbol{x}_a$ are equal to the set of all non-redundant multiplied pairs of variables $a\ in\ A = \boldsymbol{x}_a$ and $b\ in\ B = \boldsymbol{x}_h$: $A \times B = \{(a{\cdot}b) | a \in A\ and\ b \in B\}$. Together with the vector of advisor characteristics this forms a vector of variables for observed fixed effects x^f_{ah}. Since the difference between professional and lay advisors is of particular interest to us, we also interact all remaining advisor characteristics $\boldsymbol{x}_{a,-pro}$ with the professional dummy $\boldsymbol{x}_{a,pro}$:

$$\boldsymbol{x}^f_{ah} = \left[\boldsymbol{x}_a\ ,\ \boldsymbol{x}_{a,-pro}{\cdot}\boldsymbol{x}_{a,pro}\ ,\ \underbrace{\boldsymbol{x}_h{\cdot}x_{a,1}\ ,\ \boldsymbol{x}_h{\cdot}x_{a,2}\ ,\ \dots\ ,\ \boldsymbol{x}_h{\cdot}x_{a,d_a}}_{x_h \times x_a} \right]$$

Appending $\boldsymbol{x}^f_{ah}$ by an intercept and the vector of household characteristics, i.e., $[1\ ,\ \boldsymbol{x}_h\ ]$, we could estimate one large pooled regression, clustering standard errors at the advisor level to account for dependence between recommendations by the same advisor. An alternative approach first estimates the regression in (7) for each advisor independently, and then regresses estimated regression coefficients on $\boldsymbol{x}_a$ (see (8)).

$$\boldsymbol{y}_{ah} = \boldsymbol{x}_h \boldsymbol{\beta}_a + \varepsilon'_{ah} \text{ with } \varepsilon'_{ah} \sim \mathcal{N}(0, \sigma^2_{\varepsilon'}) \tag{7}$$

$$\boldsymbol{\beta}_a = \bar{\boldsymbol{\beta}} + \Delta \boldsymbol{x}_a + \boldsymbol{\zeta}_a, \ \boldsymbol{\zeta}_a \sim N(0, V_\zeta)) \tag{8}$$

The advantage of the former approach is that it delivers statistically reliable estimates of systematic links between observed advisor characteristics and investment recommendations, including of interactions between advisor and household characteristics. The advantage of the latter approach is that it measures observed heterogeneity between advisors via Δ (again including interactions) as well as unobserved heterogeneity, i.e., beyond what can be explained by $\boldsymbol{x}_a$ via $\{\zeta_a\}$ and V_ζ.

However, our design does not lend itself to this two-step approach because we obtain only a small amount of likelihood information from each advisor. Thus, independent estimates of $\boldsymbol{\beta}_a$ from each advisor's observations are extremely noisy or not even likelihood-identified. Even if we jointly estimate Equations 7 and 8 using the standard parameterization of hierarchical Bayesian models (e.g., Rossi et al. (2005)), we face an extremely high dimensional estimation problem. Both V_ζ and, especially, Δ are high-dimensional objects in our application. And the statistical reliability of estimates of Δ depends on that of V_ζ-estimates, and vice versa.

To simultaneously obtain reliable estimates of fixed effects together with reliable measures of heterogeneity, we rewrite the model as a mixed-effects model and regularize estimation of parameters corresponding to Δ and V_ζ using variable selection.

In the mixed-effects model representation the portfolio allocation rule to be estimated is, with $\boldsymbol{y}_{ah}$ representing the log risky asset share for household h recommended by advisor a:

$$\boldsymbol{y}_{ah} = \boldsymbol{x}^f_{ah} \boldsymbol{\alpha} + \boldsymbol{x}^r_h \boldsymbol{\beta}_a + \varepsilon'_{ah} \text{ with } \varepsilon'_{ah} \sim \mathcal{N}(0, \sigma^2_{\varepsilon'}), \tag{9}$$

with the following random effects distribution:

$$\begin{gathered}\boldsymbol{\beta}_a = \bar{\boldsymbol{\beta}} + \boldsymbol{\eta}_a \\ \text{with } \boldsymbol{\eta}_a \sim \mathcal{N}_{(d_h+1)}(\mathbf{0}, \boldsymbol{V}_\eta)\end{gathered}$$

The superscript in $\boldsymbol{x}^r_h$ indicates that the corresponding parameters are random effects. Also note that the random effects $\boldsymbol{\eta}_a$ are net of the variation in effects explained by $\boldsymbol{x}^f_{ah}$. Finally, parameters in the matrix Δ in (8) measuring systematic interactions between client (household) characteristics and the advisor characteristics are represented in the parameter vector $\boldsymbol{\alpha}$ in (9).

2.2. Parsimonious estimation of a linear mixed model

In this section we further develop the model parameterization for parsimonious estimation.

Parsimonious estimation of fixed effects

With d_a advisor characteristics there are $d_a{\cdot}d_r$ interaction terms added to the fixed effects coefficient vector $\boldsymbol{\alpha}$ so that $\boldsymbol{\alpha}$ contains 12*20=240 coefficients on interactions between household and advisors variables, 12 advisor characteristics, and 11 interactions of a professional dummy with all other advisor characteristics. Hence the fixed effects design matrix is

$$\boldsymbol{x}_{ah}^{f} = \left[\boldsymbol{x}_a \, , \; \boldsymbol{x}_{a,-pro}{\cdot}\boldsymbol{x}_{a,pro} \, , \; \boldsymbol{x}_h \times \boldsymbol{x}_a\right] \tag{10}$$

and $\boldsymbol{\alpha}$ contains 263 elements. The dimensionality problem has shifted from Δ to $\boldsymbol{\alpha}$. However, coefficients on the interaction terms are estimated on the pooled data over all portfolio recommendations across all advisors. Finally, we ameliorate the dimensionality problem in $\boldsymbol{\alpha}$ by introducing variable selection on $\boldsymbol{\alpha}$ as we explain next in section 2.3. Further details are available in Appendix A.

Parsimonious estimation of the random effects' covariance structure

After moving the interactions $\boldsymbol{x}_h \times \boldsymbol{x}_a$ to the fixed effects design matrix, we can rewrite our model to enable a parsimonious estimation of the random effects covariance structure as in Frühwirth-Schnatter and Tüchler (2008). First, we stack our portfolio allocation equation and then separate the mean and variation of the random effects $\boldsymbol{\beta}_a$.

1. Stacking observations by advisor

Each advisor in our sample provided risky asset share recommendations for five different household profiles. These five recommendations are stacked in a column vector $\boldsymbol{y}_a$ for each advisor. The corresponding values of the advisors characteristics and interactions (including $d_f = d_a + (d_a - 1) + d_a \cdot d_h$ variables) are stacked in the $5 \times d_f$ -dimensional fixed effects design matrix $\boldsymbol{X}_a^f$, with constant advisor characteristics within the five recommendations per advisor. Household characteristics and a leading constant are stacked in the $5 \times d_r$ dimensional random effects design matrix $\boldsymbol{X}_a^r$ with $(d_r = d_h + 1)$. Accordingly, our model for advisors $a = 1, \ldots, N$ becomes:

$$\boldsymbol{y}_a = \boldsymbol{X}_a^f \boldsymbol{\alpha} + \boldsymbol{X}_a^r \boldsymbol{\beta}_a + \boldsymbol{\varepsilon}_a, \qquad \boldsymbol{\varepsilon}_a \sim \mathcal{N}_5(\boldsymbol{0}, \sigma_\varepsilon^2 \boldsymbol{I}_5) \tag{11}$$

$$\boldsymbol{\beta}_a = \bar{\boldsymbol{\beta}} + \boldsymbol{\eta}_a, \qquad \boldsymbol{\eta}_a \sim \mathcal{N}_{d_r}(\boldsymbol{0}, \boldsymbol{V}_\eta) \tag{12}$$

2. Separating mean and variation of random effects

So far the random effects $\boldsymbol{\beta}_a$ have been normally distributed with mean $\bar{\boldsymbol{\beta}}$ and variance-covariance matrix $\boldsymbol{V}_\eta$. For the hierarchical linear mixed model as in Frühwirth-Schnatter and Tüchler (2008), we introduce $\boldsymbol{Q}$, a new covariance matrix for the random effects, which is decomposed into the lower triangular matrix Cholesky factors: $\boldsymbol{Q} = \boldsymbol{C}\boldsymbol{C}'$. The advisor specific unobserved random effects are now represented by vector $\boldsymbol{z_a}$ with length d_r which is scaled by $\boldsymbol{C}$, resulting in the following model:

$$\boldsymbol{y}_a = \boldsymbol{X}_a^f \boldsymbol{\alpha} + \boldsymbol{X}_a^r \bar{\boldsymbol{\beta}} + \boldsymbol{X}_a^r \boldsymbol{C} \boldsymbol{z_a} + \boldsymbol{\varepsilon}_a \tag{13}$$

$$\boldsymbol{\varepsilon}_a \sim \mathcal{N}_5(\boldsymbol{0}, \sigma_\varepsilon^2 \boldsymbol{I}_5) \tag{14}$$

$$\boldsymbol{z}_a \sim \mathcal{N}_5(\boldsymbol{0}, \boldsymbol{I}_{d_r}) \tag{15}$$

Both, $\boldsymbol{C}$ and $\boldsymbol{Q}$ have at most $d_r(d_r+1)/2$ non-zero elements (with d_r being the number of random effect variables in $\boldsymbol{X}_a^r$), since the off-diagonal covariance elements in $\boldsymbol{Q}$ are duplicates. Nevertheless, the number of parameters to estimate increases considerably with the number of random effects, d_r, included in the model. Since $\boldsymbol{Q}$ is required to be positive definite, parameter constraints on the elements are very complex. Therefore, Frühwirth-Schnatter and Tüchler (2008) introduce Bayesian variable selection on $\boldsymbol{C}$ instead of $\boldsymbol{Q}$. The lower triangular Cholesky factor $\boldsymbol{C}$ removes parameter constraints completely, as any arbitrary combination of element values can form a valid covariance matrix. Further, if a diagonal element of $\boldsymbol{C}$ is chosen to be zero, all other elements in the corresponding column can be set to zero. This makes it possible to quickly reduce the number of parameters to be estimated.

Consider the very simple example of $\boldsymbol{Q}$ with dimension 2×2:

$$\boldsymbol{Q} = \begin{pmatrix} 0 & 0 \\ 0 & 25 \end{pmatrix} \tag{16}$$

$\boldsymbol{Q}$ can be represented by the following two (and other) Cholesky factors, since $\boldsymbol{C}_1\boldsymbol{C}_1{}' = \boldsymbol{C}_2\boldsymbol{C}_2{}' = \boldsymbol{Q}$.

$$\boldsymbol{C}_1 = \begin{pmatrix} 0 & 0 \\ 3 & 4 \end{pmatrix}, \quad \boldsymbol{C}_2 = \begin{pmatrix} 0 & 0 \\ 0 & 5 \end{pmatrix} \tag{17}$$

If all non-zero elements are to be estimated as parameters in a model, obviously, the structure in $\boldsymbol{C}_2$ would be the smarter, more parsimonious, choice.

2.3. Adaptive parameterization within MCMC Sampler

As mentioned before, we extend Frühwirth-Schnatter and Tüchler (2008) by variable selection on the fixed effects coefficients $\boldsymbol{\alpha}$ in addition to variable selection on the slope of the random effects' covariance matrix $\boldsymbol{Q}$. Since the heterogeneity structure is a priori unknown, both, the variance-covariance matrix and the observed slope heterogeneity must be chosen to fit the data best. Variable selection is represented by two vectors of indicators, $\boldsymbol{\gamma}^f$ of length d_f for the fixed effects and $\boldsymbol{\gamma}^r$ of length $d_c = d_r(d_r+1)/2$ for random effects. While "1" represents a parameter selected into the model, a "0" indicates a parameter restricted to zero. In every iteration, the MCMC sampler cycles once through each indicator vector. With a switching probability, that depends on the prior of the indicator elements and the likelihood ratio of including versus excluding the parameter, the indicator is set from 0 to 1 or vice versa.

Constructing the design matrices W that account for selection

For the fixed effects, the binary indicator vector $\boldsymbol{\gamma}^f$, indicating which variables of the fixed effect design matrix have non-zero parameters, serves as a vector that selects the columns of the fixed effects design matrix $\boldsymbol{X}_a^f$. When selected, these columns enter the new design matrix $\boldsymbol{W}_a^f$. $\boldsymbol{W}_a^f$ is introduced to capture selection on the fixed effects design matrix $\boldsymbol{X}_a^f$. It is thus a function of $\boldsymbol{X}_a^f$ conditional on γ_f. The $5 \times q_f$ dimensional matrix $\boldsymbol{W}_a^f$ only contains columns from $\boldsymbol{X}_a^f$ that correspond to non-zero parameters in $\boldsymbol{\alpha}$. Likewise we use $\boldsymbol{\gamma}^f$ to select elements of the $\boldsymbol{\alpha}$ parameter vector to form the new coefficient vector $\boldsymbol{\alpha}^\gamma$ for the selection model.

To clarify this, consider an example with four variables in the fixed effects design matrix $\boldsymbol{X}_a^f$ and $\boldsymbol{\gamma}^f$ indicating that only the second and fourth variable should enter the model:

$$\boldsymbol{\gamma}^f = \begin{pmatrix} 0 \\ 1 \\ 0 \\ 1 \end{pmatrix} \tag{18}$$

$$\boldsymbol{\alpha} = \begin{pmatrix} 1 \\ 2 \\ 3 \\ 4 \end{pmatrix} \tag{19}$$

$$\boldsymbol{X}_a^f = \begin{pmatrix} 1 & 6 & 11 & 16 \\ 2 & 7 & 12 & 17 \\ 3 & 8 & 13 & 18 \\ 4 & 9 & 14 & 19 \\ 5 & 10 & 15 & 20 \end{pmatrix} \tag{20}$$

These result in the modified coefficient vector $\boldsymbol{\alpha}^\gamma$ and design matrix $\boldsymbol{W}_a^f$:

$$\boldsymbol{\alpha}^\gamma = \begin{pmatrix} 2 \\ 4 \end{pmatrix} \tag{21}$$

$$W_a^f = \begin{pmatrix} 6 & 16 \\ 7 & 17 \\ 8 & 18 \\ 9 & 19 \\ 10 & 20 \end{pmatrix} \tag{22}$$

We apply a similar method for the random effects. Based on the indicator vector γ^r, we first stack the selected, non-zero elements of matrix $\boldsymbol{C}$ in $\boldsymbol{C}^\gamma$. $\boldsymbol{W}_a^r$ is a function of $\boldsymbol{X}_a^r$ and $\boldsymbol{z}_a$ conditional on γ_r. We construct the new $5 \times q_r$-dimensional design matrix $\boldsymbol{W}_a^r$ by combining the random effects $\boldsymbol{z}_a$ and the original random effects design matrix $\boldsymbol{X}_a^r$ as follows: For each non-zero element c_{lm} of matrix $\boldsymbol{C}$, going from top to bottom and left to right, with l indexing rows and m columns in $\boldsymbol{C}$, the lth column of $\boldsymbol{X}_a^r$ is multiplied by random effect $z_{a,m}$ and appended to $\boldsymbol{W}_a^r$ (see also Frühwirth-Schnatter and Tüchler (2008) for details on constructing $\boldsymbol{W}_a^r$).
Consider this example with only two variables in the random effects design matrix $\boldsymbol{X}_a^r$ and only two of three elements selected from $\boldsymbol{C}$.

$$\boldsymbol{C} = \begin{pmatrix} 0 & 0 \\ 3 & 4 \end{pmatrix} \tag{23}$$

$$\boldsymbol{\gamma}^r = \begin{pmatrix} 0 \\ 1 \\ 1 \end{pmatrix} \tag{24}$$

$$\boldsymbol{X}_a^r = \begin{pmatrix} 1 & 6 \\ 2 & 7 \\ 3 & 8 \\ 4 & 9 \\ 5 & 10 \end{pmatrix} \tag{25}$$

We first look at the elements c_{lm} of $\boldsymbol{C}$ and notice that c_{21} and c_{22} are non-zero. Therefore, we know that the new design matrix $\boldsymbol{W}_a^r$ has two columns. With $l = 2$ for both elements, we take the second column of $\boldsymbol{X}_a^r$ for both cases, but multiply by $z_{a,1}$ in case of the first element and by $z_{a,2}$ in the case of the second element.

$$\boldsymbol{C}^\gamma = \begin{pmatrix} 3 \\ 4 \end{pmatrix} \tag{26}$$

$$\boldsymbol{W}_a^r = \left(\begin{pmatrix} 6 \\ 7 \\ 8 \\ 9 \\ 10 \end{pmatrix} z_{a,1} \quad \begin{pmatrix} 6 \\ 7 \\ 8 \\ 9 \\ 10 \end{pmatrix} z_{a,2} \right) \tag{27}$$

It becomes evident that forming $\boldsymbol{W}^r$ and $\boldsymbol{C}^\gamma$ serves the purpose of vectorizing matrix $\boldsymbol{C}$ by combining $\boldsymbol{X}_a^r$ and $\boldsymbol{z}_a$ after selection based on γ^r.

The MCMC sampler estimation equation

As described before, $\boldsymbol{X}_a$ includes 12 characteristics specific to advisor a stacked five times vertically (see Table A.2 for details on the variables and their scaling in the context of regression). $\boldsymbol{X}_{a,pro}$ contains only the dummy for a professional advisor, while $\boldsymbol{X}_{a,-pro}$ excludes this dummy. $\boldsymbol{X}_a^r$ is a stacked matrix of the five household profiles that advisor a examined, consisting of a constant and 20 household characteristics (see Table A.3 for details on the variables and their scaling used in the context of regression). The full estimation equation without selection includes 12 advisor characteristics, plus 11 interactions with the professional dummy, plus $20 \cdot 12 = 240$ interactions of each household characteristics with each of the 12 advisor characteristics, resulting in a coefficient vector $\boldsymbol{\alpha}$ of length 263. $\bar{\boldsymbol{\beta}}$ captures the population means of the coefficients on a constant and 20 household characteristics, giving it a length of 21. Finally, $\boldsymbol{C}$, is the lower triangular matrix Cholesky factor of dimension 21×21.

Full Estimation Model equation

$$\begin{aligned} \boldsymbol{y}_a &= \left[\boldsymbol{X}_a \;\; \boldsymbol{X}_{a,-pro} \cdot \boldsymbol{X}_{a,pro} \;\; \boldsymbol{X}_h \times \boldsymbol{X}_a \right] \boldsymbol{\alpha} + \boldsymbol{X}_a^r \bar{\boldsymbol{\beta}} + \boldsymbol{X}_a^r \boldsymbol{C} \boldsymbol{z}_a + \boldsymbol{\varepsilon}_a \\ &= \boldsymbol{X}_a^f \boldsymbol{\alpha} + \boldsymbol{X}_a^r \bar{\boldsymbol{\beta}} + \boldsymbol{X}_a^r \boldsymbol{C} \boldsymbol{z}_a + \boldsymbol{\varepsilon}_a \end{aligned} \tag{28}$$

with

$$\boldsymbol{\varepsilon}_a \sim \mathcal{N}_5(\boldsymbol{0}, \sigma_\varepsilon^2 \boldsymbol{I}_5)$$

$$\boldsymbol{z}_a \sim \mathcal{N}_5(\boldsymbol{0}, \boldsymbol{I}_{d_r})$$

In each iteration the MCMC sampler visits a possible model defined by γ^f and γ^r in proportion to its posterior support. The following estimation equation hence depends on the selection after each iteration.

Selection Model equation

$$\boldsymbol{y}_a = \boldsymbol{W}_a^f \boldsymbol{\alpha}^\gamma + \boldsymbol{X}_a^r \boldsymbol{\beta}_a + \boldsymbol{W}_a^r \boldsymbol{C}^\gamma + \boldsymbol{\varepsilon}_a \tag{29}$$

with

$$\boldsymbol{\varepsilon}_a \sim \mathcal{N}_5(\boldsymbol{0}, \sigma_\varepsilon^2 \boldsymbol{I}_5)$$

We ran the model using four MCMC-chains each with 100,000 iterations, keeping every tenth iteration and dropping 75 percent as burn-in. All four chains converge to the same posterior by visual inspection of MCMC traces. This results in 10,000 observations of each model parameter's posterior distribution. For both model types we will only report the estimates on the 21 diagonal elements of the product $Q = CC'$, i.e. the random effects' variances. Parameters for the priors represent a zero mean and standard deviation of 100 (i.e. $a_0 = b_0 = 0$; $p_{A_0} = p_{B_0} = 10.000$) and zero on all covariances for $\boldsymbol{\alpha}$, $\boldsymbol{\beta}$, $\bar{\boldsymbol{\beta}}$ and $\boldsymbol{C}$. To facilitate interpretability we mean-center all variables. The reason for this is that, when adding an interaction term of two variables to the regression, the coefficient on the lower order term of the first variable does not represent an effect at the population average, it represents the effect conditional on the second variable. Without centering, this is conditional to the second variable being equal to zero. With centered variables, estimates for lower order terms are conditional on the population mean (see Dalal and Zickar (2012)), while the estimate for the interaction term is not affected by centering. A detailed description of the MCMC sampler, the prior choice and the results of testing the sampler are in Appendix A.

3. What theory predicts: Determinants of risky asset allocations

The main goals of our analyses are the identification of the determinants of portfolio recommendations and the assessment and comparison of heterogeneity in recommended risky asset shares for lay and professional advisors. There are two ways in which the individual opinions and preferences of advisors can influence their portfolio recommendations. Advisors might have an opinion about a baseline risky asset share, regardless of their clients' characteristics, i.e. the level might shift depending on their expectations about returns and risk or their individual risk preferences. In addition, advisors might differ in terms of how they map client characteristics into risky asset share recommendations, e.g., by how much they would lower risk if a client was approaching retirement. In this section, we summarize normative predictions from portfolio theory related to this mapping.

In Table 4 we provide an overview of important determinants of optimal risky asset allocations as provided by normative theory and observational studies. The first three factors stem from the seminal portfolio allocation rule (see section 2.1). Samuelson and Merton pioneered optimal portfolio allocation models that optimized consumption vs. saving over multiple periods. While higher expected returns and preferences for risk obviously increase the optimal risky asset share, the impact of **financial wealth** on risk taking is more complex. At first glance, richer investors should have more risk-taking potential, the risky asset share would increase in the amount invested. But the impact of financial wealth on the optimal allocation strongly depends on other factors. First, we know that younger and older investors are prescribed higher and lower risky asset shares respectively because of differences in their investment horizon. This is undisputed and implies that, over the life cycle, the risky asset share on financial wealth decreases. Now, if we assume that risk preferences are constant for an individual investor and, more specifically, that the investor has constant relative risk aversion (CRRA), an increase in total wealth would lead to higher absolute risk taking, whereas the preferred share of capital invested into risky assets, i.e., relative risk, remains constant. Early in life, an investor has no financial savings and holds only human capital, which is equal to her discounted labor income over the life span. Total wealth is equal to human capital + financial wealth. But an investor is not able to invest in risky assets if she holds no liquid financial wealth. This is especially common early in life and still common among middle-agers who hold tangible assets only in form of illiquid housing equity. Therefore, to reach a preferred adequate share of risky assets early in life, the risky asset share on financial wealth will stay at 100% until the risky share on total wealth has reached its desired level. Any increase in financial wealth, up until this point, will increase risk taking. From this point on, an increase in financial wealth would lead to a decrease in the risky asset share on financial wealth, keeping risk taking on total wealth constant. This holds for as long as wealth other than financial wealth (e.g. human capital or housing equity) does not reach a marginally small and unimportant share of total wealth. If there is practically only financial wealth, increases in financial wealth would c.p. only lead to an increase in absolute risk taking, but the risky asset share should remain constant, given that we are considering an individual with constant relative risk aversion. In summary, the c.p. effect of changes in financial wealth on relative risk taking depends on the ratio of financial wealth to total wealth. In comparison, higher levels of **labor income** c.p. increase total wealth, not financial wealth, which unambiguously leads to a higher risk taking potential (see e.g. Bodie et al. (1992), Cocco et al. (2005)).

This decreasing effect of financial wealth on risk taking when financial wealth becomes the dominant form of capital due to the diminishing role of human capital over the life cycle, as elaborated above, is counteracted by an additional rationale. Private investors are usually consumers with a certain **subsistence** or consumption **habit** level. To avoid consumption below subsistence (i.e. starving), investors should optimally invest a sufficient amount of their financial wealth in strictly riskless assets to generate a buffer stock, for example to bridge periods of unemployment. As soon as savings are high enough to guarantee subsistence for the foreseeable future, an additional increase in financial wealth finally increases risk taking potential, conditional on a certain life cycle period (see e.g. Gomes and Michaelides (2003), Polkovnichenko (2007)). In conclusion, the ceteris paribus effect of financial wealth is ambiguous: It depends on the level and even interactions of other household characteristics.

As mentioned above, **age** primarily determines the investment horizon for retirement savings but is clearly related to life-cycle effects. Empirically, it has remained a challenging or rather impossible task to measure how portfolios vary c.p. across age groups. As Ameriks and Zeldes (2004) state: "A typical rule of thumb [of financial advisors] is that the percentage of an investor's portfolio of financial assets that is held in equities should equal 100 minus her age." But checking this prescription empirically is impossible because researchers cannot disentangle age from cohort effects. Cocco et al. (2005) show in a range of life-cycle portfolio optimizations that the risky asset portfolio share should be 100% early in life and decline over time until retirement, after which the share should be relatively stable or even slightly increasing.

Finally, we consider the impact of **real estate** or housing wealth on risk taking. Expanding on the analysis of Cocco (2005), who argues that house price risk crowds out stock holdings, Chetty and Szeidl (2012) emphasize the importance of distinguishing the value of housing as an asset from the associated mortgage loan. In their portfolio model, housing enters the utility function to model the relative importance of housing and consumption. Chetty and Szeidl show that an exogenous increase in house prices increases home equity wealth and therefore also total wealth. This induces CRRA households to adjust their risky asset portfolio share in the financial portfolio to keep a constant fraction of total wealth invested in the risky asset. This effect is countered by several channels. First, an increase in house prices increases the risk exposure in total wealth as house prices fluctuate. Second, when house prices and stock returns covary, increases in housing wealth will induce negative hedging demand for the risky asset. Finally, higher housing wealth can also be financed entirely by mortgage debt, which reduces life-time wealth when the interest on a mortgage exceeds the risk-free rate. In this case, the risky asset share decreases unambiguously. In summary, it is clear that housing wealth is subject to price changes and therefore risky. As indicated by Cocco, households with low wealth will be concerned with reducing mortgage debt instead of further increasing risk by buying stocks. But the tendency when considering housing equity and debt separately is clear. More housing equity increases total wealth and thus risk taking potential. More housing debt decreases total wealth, and increases risk on total wealth, therefore decreasing the risk taking potential related to financial wealth.

Table 4.
Normative determinants of the optimal risky asset allocation
The table shows a summary of important determinants for portfolio allocations identified in the normative literature on portfolio choice. The second column reports how the factors enter our estimation model. Adv. baseline refers to the risky asset share recommended by advisors irrespective of client characteristics. Return expectations (in the first row of the table below) were elicited from participants and not provided in household profiles. The impact of all other factors can be observed in the estimated portfolio rules, i.e. the mapping of household variables into risky shares (HH mapping rule). The third column illustrates the direction of the expected effect and the last column provides some key references for each factor.

Factor in portfolio theory	**Entering via**	**Normative /predicted impact**	**Framework**	**Main References**
E(Asset return)	Adv. baseline	↗	Expected utility theory	Huang and Litzenberger (1988)
Risk tolerance	HH mapping rule	↗	Optimal risk allocation	Samuelson (1969)
(Financial) Wealth	HH mapping rule	↗↘	Life cycle consumption/saving	Merton (1969)
Age	HH mapping rule	↘	Life cycle models	Ameriks and Zeldes (2004), Cocco et al. (2005)
Labor income / human capital	HH mapping rule	↗	Labor flexibility	Bodie et al. (1992), Cocco et al. (2005)
Income risk	HH mapping rule	↘	Background risks (self-employm.)	Heaton and Lucas (2000b)
Habit / subsistence level	HH mapping rule	↘	Life cycle models with habit	Gomes and Michaelides (2003), Polkovnichenko (2007)
Real estate wealth	HH mapping rule	↗	Housing equity vs. debt	Chetty et al. (2017)
Real estate debt	HH mapping rule	↘		

4. Results

We now assess what our model's posterior estimates reveal about the determinants of portfolio recommendations, compare lay and portfolio advisors regarding their average recommendations and the heterogeneity of these recommendations. This section presents results based on our model's coefficient estimates, in section 4.1, and the results of using the model to predict risky asset share recommendations, in section 4.2.

4.1. Estimation results

Determinants of the advisors' portfolio allocation rules: *Which household characteristics appear to be important factors in the context of mapping risky asset share recommendations?*

Table 6 Panel C reports the coefficients on household variables in $\bar{\boldsymbol{\beta}}$ representing the random effects' means, i.e. the portfolio rule of an average advisor on the average household. It appears that many household characteristics are used to determine the optimal risky asset share. To gain some intuition, we begin by exploring the effect size of important household characteristics. For example, the effect of being self-employed or reporting a risk tolerance that is one unit larger (on a qualitative scale from zero to five) is around 0.2 in absolute terms (the exact coefficients are 0.220 and -0.212 respectively). Adding 0.2 to the risky share on the log scale would increase a share of 23%[9] by roughly 20% or five percentage points. For income, entering the regression in logs, the coefficient of 0.253 means that a 100% increase in income would result in a 25% rise in the risky asset share. Hence, a 25% increase in income would fully counteract the risk of being self-employed or reporting a one-unit-lower risk tolerance.

Concerning the mapping of age, the average investor at the age of 45.5 would need to be 8 years older to see a decrease of around 5% in the recommended risky asset share; see Figure 2. This corresponds to having almost seven more kids in the household or a large decrease of around 70 percentage points in the ratio of safe income. Both are not realistic even if the corresponding averages of 1.6 and 85.7% were much lower. In comparison, the effects of kids and safe income are thus much smaller. Other variables suggested by theory appear to be regarded as less important by our participants: The insignificant coefficient on financial wealth, i.e. the investment amount, is more than ten times smaller, while the coefficients on housing wealth and debt are 25 and 18 times smaller than the income coefficient, but both are significant.

Comparing the significant coefficients to the summary on normative prescriptions in Table 4, we see that all factors enter the portfolio rule as predicted. The importance of labor income mirrors the outstanding role of human capital in the literature on portfolio allocation (see e.g. Bodie et al. (1992)). As expected, age enters the portfolio rule as a salient factor, given that it has been reported to be an important heuristic factor for financial advisors, as discussed in the previous section. The effect of financial wealth, represented by the investment amount in our experiment, is positive and thus in line with the empirical evidence of Calvet and Sodini (2014), who find that

9. For the average advisor and the average household profile the risky asset share is equal to $exp(constant)$, the constant is the first element in $\bar{\boldsymbol{\beta}}$: $exp(-1.474) \approx 23\%$

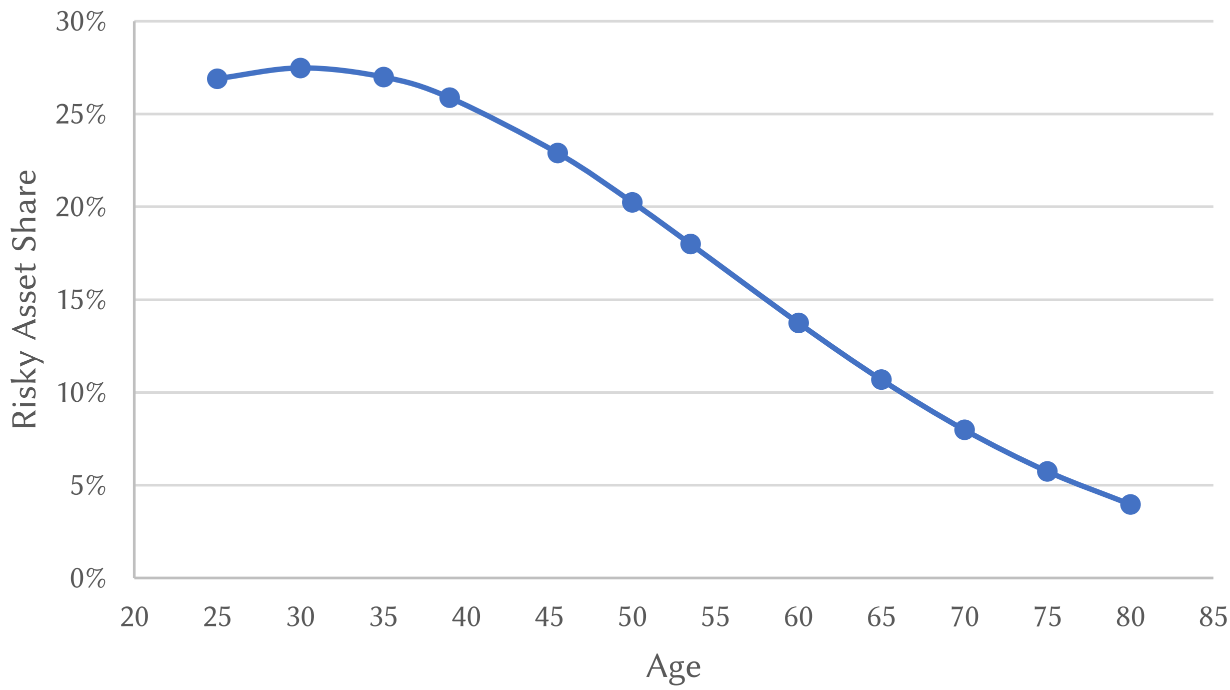


Figure 2.
Estimated age profile of the recommended risky asset share
The curve is based on the coefficients on *age* and *age*^{2} in $\bar{\boldsymbol{\beta}}$ and thus represent the c.p. change in the risky asset share due to an age variation for the average household profile and average advisor.

risk taking increases with financial wealth. This contrasts with the theory on the ratio between financial and human capital, which predicts an increasing risk aversion with growing financial wealth. However, the small magnitude and insignificance of the effect indicates that the direction is not entirely clear.

During the experiments, we asked the participants to mark the household characteristics they would use to determine the risky asset share. Table 5 shows the percentage of participants that selected each characteristic. The dominant determinants identified in our model appear in similar order in the check-box ranking collected during the experiments. Both advisor groups show a similar order, while lay advisors appear slightly less motivated to select characteristics, with only 77% marking any characteristic compared with 82% in the case of professional advisors.

In summary, there is evidence that risk tolerance, age, and income are the most important factors in advisors' portfolio rules. Wealth, whether financial or real estate, only plays a secondary role when judging by effect size.

Table 5.
Profile characteristics check-boxed in experiment
The table reports the advisor group-specific fraction of participants that checked the corresponding household characteristic (in rows) during the portfolio experiment. The last line states the fraction of participants that checked any box.

	Prof. Adv.	Lay Adv.
Age	0.75	0.51
Income	0.71	0.56
Risk Tolerance	0.69	0.44
Investment Amount	0.64	0.47
Investment Exper.	0.60	0.35
Safe Income	0.54	0.44
Real Estate Wealth	0.48	0.30
Self Employed	0.46	0.40
Real Estate Debt	0.44	0.26
College	0.42	0.27
Food Expenditure	0.39	0.27
Married	0.37	0.24
Kids	0.27	0.20
Male	0.08	0.09
Checked Any	0.82	0.77

Differences in average allocations - intercept heterogeneity: *What factors influence the heterogeneity in recommendations?*

We start assessing the impact of advisor characteristics on the risky asset share by examining the corresponding coefficients of the Estimation Model; see Panel A of Table 6. The first column shows the selection ratio, i.e. the percentage of iterations of the MCMC-sampler with non-zero coefficient estimates. The last three columns report statistics on the posterior distribution. The first set of rows contains the advisor characteristics and the second set of rows the interactions of the professional dummy with all other advisor characteristics. We see that the advisors' age, own allocation and long-term return expectations have a positive and significant impact on the recommended risky asset share. Interestingly, short-term return expectations (the coefficient on "Ret.Exp.1yr"), risk tolerance and its interaction with the professional advisor dummy report selection rates above 90%, indicating their importance; however, the estimates are not significant at the 10% level. Since we estimate effects at the population average, doubling the risky asset share of an advisor's own private portfolio from the mean, i.e. around 40% to 80%, would lead c.p. to an increase of around 6.6% in the average recommended risky asset share. By comparison, an increase in the advisor's age by ten years (e.g. from 45 to 55) is associated with an increase of 9% in the recommended asset share. The variable for return expectations over ten years represents the advisors' belief, in form of a percentage probability, that the DAX return will, on average over ten years, exceed 10% annually. Given the posterior mean of 0.004, an 8% higher risky asset share would thus require a 20 percentage points increase in the belief, e.g. an increase from 20% to 40%. Notably, for each of the three advisor characteristics, the change needed to trigger a comparable (the same order of magnitude) increase in the risky asset share recommendation is arguably close to one standard deviation estimated across the overall advisor population (see the last column in Table 3). They appear to be of similar importance in their role as determinants of the measured recommendations.

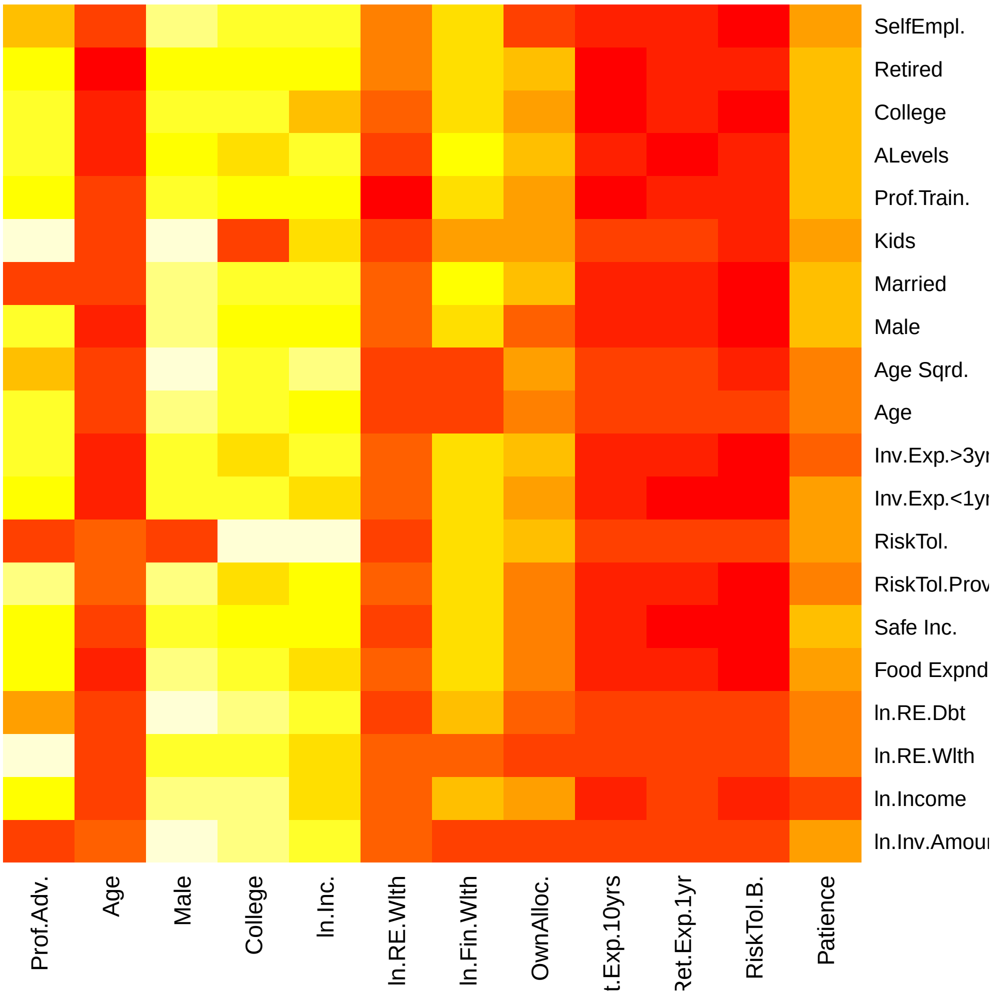


Figure 3.
MCMC sampler selection ratio
Heatmap on the advisor-household interaction coefficients. Higher selection ratios are redder, lower ratios whiter. For example, age squared × 10 year return expectations achieved the maximum of 100% selection. Safe Income × Male received the minimum with 2.16% selection.

Differences in the mapping of household characteristics into optimal portfolio allocations - slope heterogeneity: *What observable advisor characteristics cause heterogeneity in portfolio rules?*

Panel B of Table 6 reports the coefficients of interactions between advisor and household characteristics and helps us to further identify sources of heterogeneity. Heterogeneity can stem either from household characteristics that are not homogeneously mapped into recommendations or from certain advisor groups, identified by their demographic characteristics that vary in their optimal portfolio mapping rules. Surprisingly, few interactions appear to be significant. For example, male advisors seem to react much less sensitively to variations in risk tolerance given the posterior mean of -0.109 on the $Risk\ Tolerance \cdot Male$ interaction.

In general, many significant interactions in columns would suggest advisor characteristics that identify groups of advisors with differences in the allocation rule, whereas many significant interactions in rows would identify household characteristics that are heterogeneously mapped into the risky asset share across advisors. By this logic, two candidates for heterogeneity within the set of observed advisor characteristics are the dummy for professional advisors and the advisors' risk tolerance; see Panel B of Table 6.

The most salient household characteristic in terms of observable heterogeneity is the investment amount, i.e. households' financial wealth. To verify these findings, we exploit the selection mechanism of our Bayesian MCMC sampler and examine the frequency of selecting interactions into the model, i.e. the rate at which coefficients were selected to be non-zero over the simulation iterations, when running the sampler. In Figure 3, we illustrate the selection ratio in a heatmap, with advisor characteristics in columns and household characteristics in rows. Selection rates range from zero (white) to 100% (red). The visualization clearly identifies the advisors' age, real estate wealth, return expectations and risk tolerance as determinants of heterogeneity in allocation rules. Interestingly, the professional dummy only shows high selection rates on interactions with a small set of household characteristics such as the investment amount (i.e. financial wealth), real estate debt, and risk tolerance. Scanning the rows, we see that households' investment amount, real estate wealth and debt as well as age might be factors that enter advisors' portfolio rules more heterogeneously.

Differences in recommendations between professional and lay advisors: *Do professional advisors use different portfolio rules?*

We again analyze Panel A of Table 6 to investigate if professional advisors show differences in risky asset share recommendations without taking into account differences in the mapping of household characteristics by first taking a look at intercept heterogeneity. In column (2) of Panel A, which states the posterior mean, we see that age and the advisors' own allocation are significant determinants of the average recommended risky asset share. The interactions between the professional advisor dummy and the remaining advisor characteristics (in the second set of rows) are insignificant, but the selection rate (column (1)) reveals additional factors that suggest differences between advisor groups. The interactions on age, return expectations and risk tolerance all show selection rates above 80%. However, as identified by Foerster et al. (2017), the dominant role of advisors' own portfolio allocation in their recommendations, in particular, shows that there is no significant difference between professional and lay advisors. Finally, to analyze heterogeneity in mapping rules, i.e. slope heterogeneity between lay and professional advisors, we look at column (1) in Panel B. Differences in the allocations rules between lay and professional advisors are significant for households' investment amount and risk tolerance. As outlined above in the heatmap, i.e. Figure 3, we see that both factors are selected at a rate of close to 100%. Households' investment amount and risk tolerance enter the portfolio allocation rule positively across advisors (see column (2) in Panel C of Table 6) but only risk tolerance is statistically significant. Nevertheless, professional advisors increase the recommended risky asset share more strongly after c.p. increases in both households' investment amount or risk tolerance. We recall that all coefficients are estimated after mean-centering the regression variables, and that around half of our participant sample is made up of professional advisors. The coefficient of 0.157 on the $Prof.Adv.{\cdot}Inv.Amount$ interaction thus means that the investment amount shows a positive elasticity of around 0.10 (= 0.022 + 0.5•0.157) in the case of professional advisors and a negative elasticity of around -0.057 (= 0.022 − 0.5•0.157) in the case of lay advisors. A doubling of the investment amount would thus approximately result in an 10% higher (around 6% lower) risky asset share recommendation for professional (lay) advisors. Selection rates also suggest a stronger reaction to real estate debt and being married among professional advisors. Overall, the evidence suggests that professional advisors implement normative prescriptions more seriously, i.e. with more weight on important characteristics, than lay advisors.

Table 6.
The Estimation Model's posterior estimates
The table shows summary statistics on the posterior distributions of our model coefficients. The model includes advisor characteristics and interactions of a binary indicator for professional advisors with the remaining characteristics (Panel A), household profile variables (Panel C) and their interactions with each advisor characteristic (Panel B). All variables are centered at the population mean. The first column in the table below reports the fraction of MCMC iterations for which a coefficient was selected into the model, i.e. where the coefficient was non-zero. The second column reports posterior means, columns three and four the 5th and 95th quantiles of the posterior distribution for each coefficient estimate. Panel B only includes posterior means due to space constraints. Panel C includes no column for selection, as our MCMC sampler introduces selection only on fixed effects. We ran the model using four MCMC-chains with 100.000 iterations each, keeping every tenth iteration and dropping 75 percent as burn-in. This results in 10,000 observations of each model parameter's posterior distribution. In all panels, a star (*) represents a p-value of at least 10%, which means that the 5 to 95 percent posterior interval is either above or below zero.

Panel A: Advisor fixed effects in α

	Selection	Posterior statistics		
	Gamma FE	Mean	5% Quantile	95% Quantile
Prof.Adv.	82.42%	−0.154	−0.308	0.000
Age	99.66%	0.009*	0.003	0.014
Male	21.95%	−0.010	−0.111	0.019
College	25.8 %	−0.008	−0.090	0.029
Income	31.69%	−0.012	−0.102	0.023
RE.Wlth	76.28%	−0.002	−0.013	0.009
Fin.Wlth	47.72%	−0.009	−0.054	0.019
OwnAlloc.	99.96%	0.066*	0.038	0.094
Ret.Exp.10yrs	98.94%	0.004*	0.001	0.007
Ret.Exp.1yr	91.78%	0.001	−0.002	0.004
RiskTol.B.	97.43%	0.002	0.000	0.004
Patience	71.45%	−0.014	−0.043	0.004
Prof.Adv.				
• Age	85.2 %	−0.004	−0.014	0.003
• Male	11.79%	0.004	0.000	0.065
• College	15.23%	0.009	0.000	0.113
• Income	49.17%	−0.096	−0.328	0.000
• RE.Wlth	64.98%	−0.002	−0.022	0.015
• Fin.Wlth	29.76%	0.001	−0.046	0.058
• OwnAlloc.	45.44%	0.010	−0.017	0.062
• Ret.Exp.10yrs	88.62%	−0.002	−0.008	0.003
• Ret.Exp.1yr	89.44%	−0.002	−0.008	0.003
• RiskTol.B.	91.27%	0.001	−0.002	0.005
• Patience	52.44%	−0.014	−0.064	0.012

Table 6.
-Continued
The table shows the posterior means of coefficients on the interactions of household profile variables (rows) with advisor characteristics (columns). A star * represents a p-value of at least 10%, i.e. the 5 to 95 percent posterior interval is either above or below zero.

Panel B: Advisor characteristic × household profile interactions in α

Advisor × Profile	Prof. Adv.	Age	Male	College	Income	Hous. Wlth	Fin. Wlth	Curr. Alloc.	Ret.Exp. 10yrs	Ret.Exp. 1yr	Risk Toler.	Patience
Inv.Amount	0.157*	0.000	0.004	−0.002	0.012	−0.003	0.037	−0.020*	0.001	0.001	0.000	−0.005
Income	−0.051	0.000	0.007	−0.010	0.043	0.005	−0.021	−0.007	−0.002	0.000	0.003	0.029
RE.Wlth	−0.003	0.000	0.009	0.006	−0.007	0.001	0.006	−0.004*	0.000	0.000	0.000	−0.001
RE.Dbt	−0.014	0.000	−0.001	0.001	0.001	−0.001	0.000	−0.003	0.000	0.000	0.000	0.000
Food Expnd.	0.025	0.001	−0.012	−0.010	0.034	0.002	−0.004	0.010	0.000	0.001	−0.001	0.007
Safe Inc.	−0.033	0.004	−0.002	−0.036	0.006	0.021	0.007	−0.046	−0.004	−0.013*	0.011*	0.013
RiskTol.Prov.	−0.019	0.001	0.014	0.083	0.049	0.012	−0.023	0.033	−0.005	−0.005	−0.005	0.032
RiskTol.	0.157*	0.000	−0.109*	−0.009	−0.001	−0.006	0.009	−0.004	0.001	0.001	0.001	−0.005
Inv.Exp.<1yr	0.023	−0.001	0.001	0.012	−0.029	0.004	0.000	−0.001	0.001	0.003	0.002	−0.004
Inv.Exp.>3yrs	0.025	−0.004	−0.011	−0.041	0.009	0.002	−0.008	0.002	0.000	−0.001	0.004*	−0.021
Age	0.010	0.000	0.003	0.002	0.003	−0.003	0.016	0.002	0.001	−0.001	0.000	0.003
Age Sqrd.	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
Male	0.023	0.006	−0.003	−0.025	−0.029	−0.002	−0.008	−0.024	−0.002	0.002	0.003*	0.004
Married	−0.198	0.002	0.009	0.018	−0.011	−0.005	−0.001	0.005	−0.001	−0.001	0.005*	0.006
Kids	0.006	0.000	0.008	−0.076	0.022	−0.003	−0.011	0.002	0.000	0.000	0.001	0.004
Prof.Train.	−0.017	0.001	−0.016	−0.017	0.004	0.021	0.002	0.008	−0.004	0.000	−0.001	−0.002
ALevels	−0.022	−0.002	0.061	0.064	−0.014	0.012	0.001	0.001	0.001	−0.005	0.000	0.003
College	0.018	0.002	−0.018	−0.014	−0.062	0.006	−0.007	0.000	−0.002	0.001	0.003	−0.001
Retired	0.020	0.013	−0.006	0.002	−0.006	0.008	−0.019	0.010	0.010	0.001	−0.001	0.002
SelfEmpl.	−0.118	0.001	−0.002	0.027	0.009	−0.005	−0.008	0.048	0.000	−0.003	0.003	0.017

Table 6.
-Continued

Panel C: Coefficients on household profiles and constant - mean of random effects $\bar{\boldsymbol{\beta}}$			
	Mean	5% Quantile	95% Quantile
Const	−1.474*	−1.536	−1.411
ln.Inv.Amount	0.022	−0.009	0.054
ln.Income	0.253*	0.186	0.319
ln.RE.Wlth	0.010*	0.004	0.017
ln.RE.Dbt	−0.014*	−0.021	−0.006
Food Expnd.	0.008	−0.049	0.064
Safe Inc.	0.281*	0.095	0.468
RiskTol.Prov.	−0.522*	−0.651	−0.391
RiskTol.	0.220*	0.189	0.249
Inv.Exp.<1yr	−0.057	−0.136	0.021
Inv.Exp.>3yrs	0.028	−0.051	0.107
Age	0.047*	0.029	0.066
Age Sqrd.	−0.001*	−0.001	−0.001
Male	−0.040	−0.106	0.026
Married	−0.038	−0.116	0.039
Kids	−0.031*	−0.056	−0.005
Prof.Train.	−0.008	−0.100	0.086
ALevels	−0.044	−0.135	0.045
College	0.032	−0.060	0.122
Retired	−0.122	−0.339	0.090
SelfEmpl.	−0.212*	−0.328	−0.097

4.2. Prediction Results

To further investigate the determinants of the risky asset share recommendations, their heterogeneity, and differences between the two advisor groups, we generate predicted recommendations for each household profile and advisor combination in the sample using our estimation model (see equation (28) and Table 6 for coefficient estimates). In our experiment, each participant was presented with five household profiles. With 871 participants, we have $4,355$ household profiles in total. For each of these household profiles, we will predict the recommended risky asset shares for each of the 871 advisors, resulting in $4,355 \cdot 871 = 3,793,205$ pairings and corresponding predicted risky asset shares. For each household profile / advisor paring we use the corresponding values of household and advisor variables to generate estimates for each draw of the MCMC sampler. Predictions are expected, i.e. mean, recommendations over the $10,000$ iterations kept from the MCMC chains after burn-in.

For illustration, we have sorted the household profiles by a) the average portfolio recommendation and b) their heterogeneity (standard deviation) in portfolio recommendations across professional and lay advisors separately. The sorting helps us identify household characteristics that influence portfolio recommendations. We calculate means of household profile characteristics within quartiles of the recommendation averages and heterogeneity (standard deviations). In the first two quartiles, lay advisors show higher average recommendations and higher heterogeneity. We confirm that, overall, average recommendations are larger for lay advisors who also tend to give more heterogeneous recommendations; see regression Table 9 for the corresponding t-tests. The average recommendation is 66.8% for lay advisors, whereas professional advisors recommend a risky asset share which is 2.3 percentage points below that. Despite the fact that quartile means range from 42% to 87% for professional advisors and from only 52% to 81% for lay advisors (see Table 7), lay advisors' recommendations are estimated to be more volatile than those of professional advisors (see column (4) in Table 9).

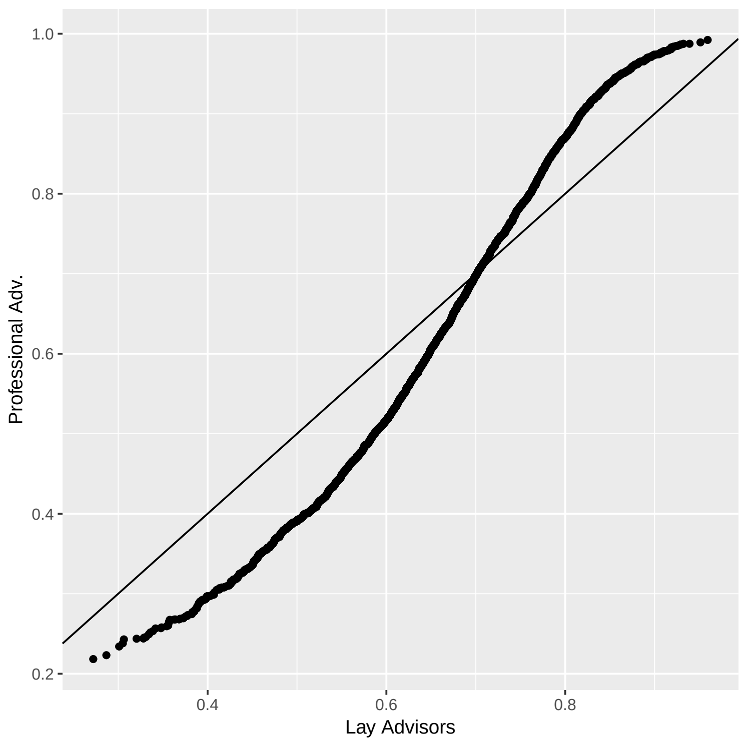


Figure 4.
Average predicted risky asset share recommendations by advisor group
Each dot represents one of the 4355 household profiles. The y-axis represents the average risky asset share recommendation across professional advisors, the x-axis the average across lay advisors. The straight line represents all points at which the average recommendation of both groups coincide.

Average recommendations

Figure 4 illustrates how close the averages of predicted recommendations are for professional and lay advisors. The straight 45 degree line covers all points where group averages would coincide. Each dot represents a household profile; the corresponding mean allocation by lay advisors is on the x-axis and the mean allocation by professional advisors on the y-axis. We see that lay advisors are likely to recommend more risky allocations, whereas their professional counterparts recommend less risky portfolios. This relationship flips for households with recommendations of a risky asset share of above 70%, i.e. the point where dots cross the parity line. Recommendations above 70% are lower for lay advisors than for professional ones. A reason for this could be the implementation of simple heuristics in the case of lay advisors. Instead of giving extreme recommendations like 0% or 100%, which are hard to justify, they might gravitate towards the mean or even a 50/50 allocation.

We now want to compare household profiles that receive different risky asset share recommendations; see Table 7 that reports summary statistics on household profiles by risky asset share quartiles. We recall from the results in section 4.1 that the households' investment amount, income, risk tolerance, and age are among the most important determinants of the risky asset share. We have seen that significant differences in the mapping rules of professional and lay advisors exist in households' investment amount and risk tolerance; see column (1), Table 6, Panel B. We illustrate differences between advisor groups across the defined quartiles in these selected household characteristics. In Figure 5, we see that lay advisors' average recommendations increase for household profiles with higher income and risk tolerance, while they decrease in terms of investment amount and age. While professional advisors' predicted recommendations show a stronger, but equally directed reaction to risk tolerance and age, a higher risk allocation is more frequently recommended to investors with higher investment amounts. The differences in the levels of risk tolerance and age are not surprising. Both variables are standard indicators of clients' risk-taking potential regarding financial advice. The reaction to the investment amount on the other hand is interesting. It might be that lay advisors wish to protect larger amounts of wealth from the high volatility on the stock markets, whereas professional advisors seek (high) returns. This argument is supported by the vast amounts of cash that Germans save in accounts with no or only small interest.

In Table 10 we show exemplary household profiles that receive either high (column (1)) or low (column (4)) average recommendations from lay or professional advisors, and household profiles that receive the most contradictory recommendations from the two groups (columns (2) and (3)). In column (1), we see that lay and professional advisors agree on recommending a high risky asset share for a young investor with very high housing equity, high income and low financial wealth. This is perfectly consistent with the normative theory discussed in the previous section. For the given profile, financial wealth represents only a small fraction of total wealth, which also includes housing equity and human capital, resulting in a very high risk taking potential. The profile in column (4) is another interesting example. Both advisor groups would recommend a risky share of around 30% although housing equity is moderate and the investment amount is large. However, the household is marked as retired with only a small annual income, justifying a smaller risky asset share. Columns (2) and (3) reveal that professional and lay advisors do not agree on household profiles that are more complicated and deliver mixed signals. For example, in column (3) the investor is young and wealthy but receives only a small labor income; this results in a high recommendation from professional advisors and a low recommendation from lay advisors, which might indicate that lay advisors are very sensitive when it comes to regular income.

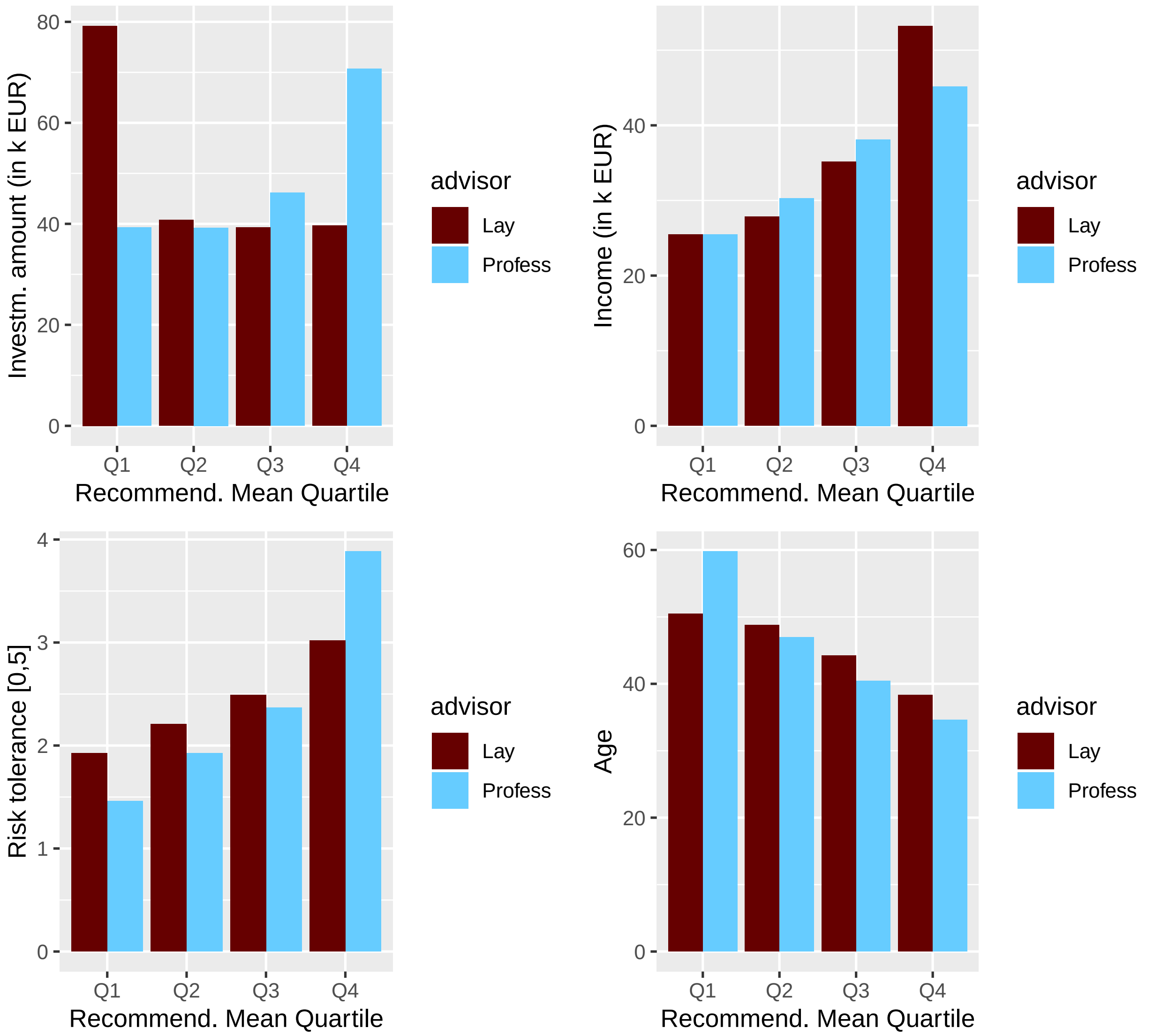


Figure 5.
Averages of household profile characteristics by quartiles of recommendation mean
Household profiles are sorted into quartiles of mean recommendations across lay and professional advisors, represented on the x-axis. The y-axis shows the mean in the corresponding household characteristic for each quartile.

Heterogeneity in recommendations

After sorting household profiles by the heterogeneity in advisors' recommendations (estimated standard deviation), surprisingly, both advisor groups appear to be quite similar. In Figure 6 (and Table 8), we observe that higher heterogeneity in recommendations (i.e. a higher quartile) is associated with household profiles that show lower investment amounts, lower income, and lower risk tolerance. Whereas age is higher in household profiles that face higher recommendation uncertainty.

It seems that poorer, more risk averse, and older investors face a larger variety of recommendations on how risky their investments should be. It is plausible that larger financial wealth increases the opportunity cost of avoiding market participation and advice becomes more congruent. Furthermore, one could argue that the financially poorer should save more and invest in assets with large returns to escape poverty in the long run, but conversely, one could argue that poorer households should invest less risky, e.g. to sustain consumption.

In Table 11 column (1) we see that recommendation heterogeneity is large for a household close to retirement with only low income and small financial and housing wealth. A household profile that receives the most homogeneous recommendation from both advisor groups is shown in column (4). The investor is very young, has very large housing equity and income, but only a small investment amount. The finding that old age is associated with higher heterogeneity seems counter-intuitive. Normative theory clearly prescribes to decrease risk as the investment horizon shortens over the life-cycle. The pattern is more pronounced for lay advisors (see the bar chart on age in Figure 6), maybe because they are less acquainted with standard investment advice and the implications of varying investment horizons, which are basic concepts for any professional advisor. Heterogeneity might arise due to different interpretations of the data. There are certainly well situated elders that are asset-rich and receive a sufficient amount of pension to cover all expenses. If substantial amounts of wealth exist that are not planned to be used as retirement income, it would certainly make sense to invest more riskily even after retirement to prevent devaluation of financial wealth due to inflation. For a household profile that is not unambiguously mapped into an optimal risky asset allocation, many scenarios about the households living circumstances and thus optimal allocations could arise that are plausible given the provided demographics.

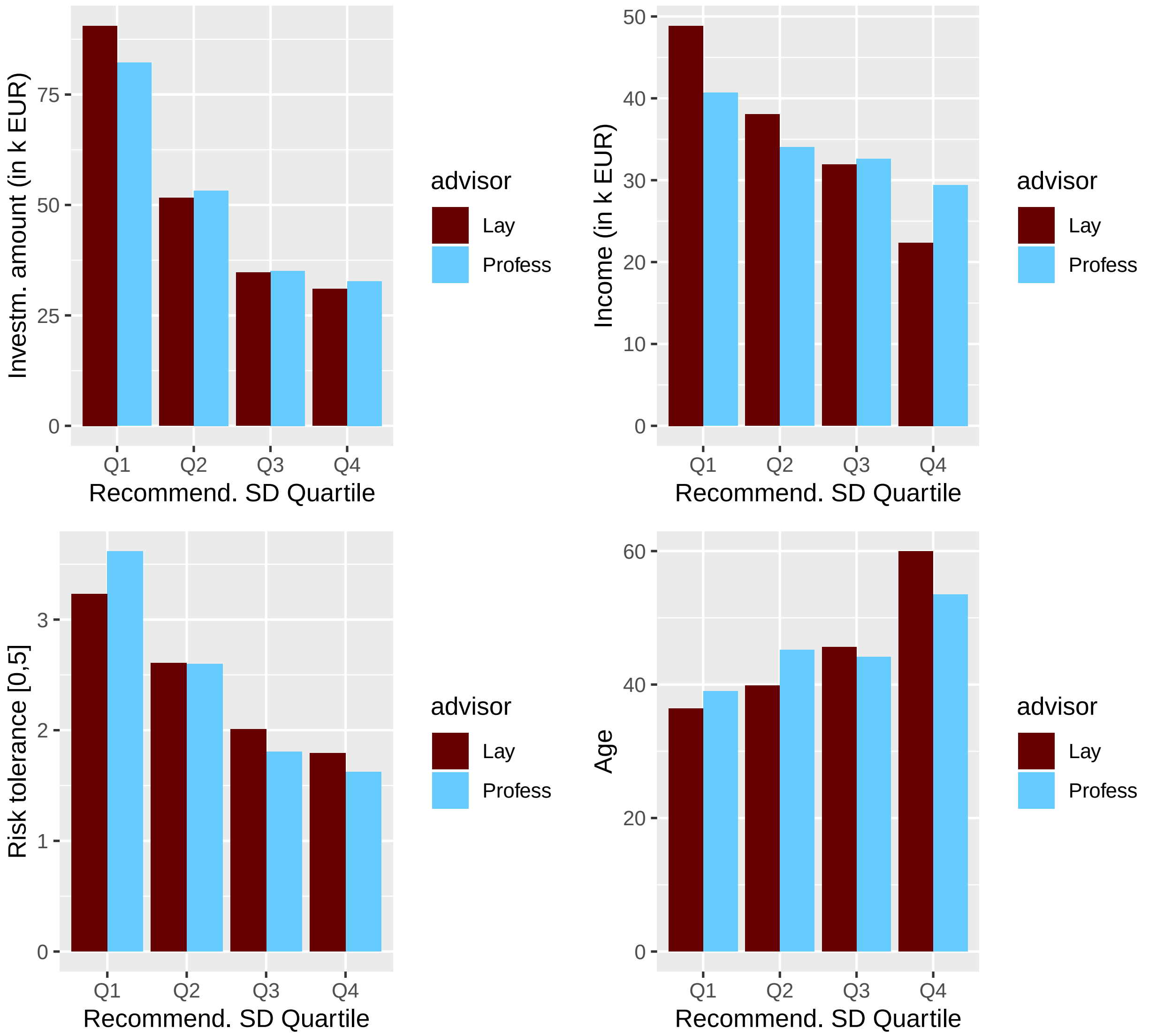


Figure 6.
Averages of household profile characteristics by quartiles of recommendation heterogeneity
Household profiles are sorted into quartiles of the standard deviation (SD) of recommendations across lay and professional advisors, represented on the x-axis. The y-axis shows the mean in the corresponding household characteristic for each quartile.

Table 7.
Household profiles by quartiles of the risky asset shares (recommendation averages)
Household profiles are sorted into quartiles of the average recommended risky asset share across lay and professional advisors, represented by columns. The first row contains the quartile means of recommendation averages, the following rows report household profile characteristics averaged within the corresponding quartile. For example, the quartile of households with the highest risky asset share recommendation by lay advisors is shown in column (4).

	Lay Advisors				Professional Advisors			
	Q1	Q2	Q3	Q4	Q1	Q2	Q3	Q4
Risky Share Average	52.2	63.5	70.9	80.6	42.0	57.8	71.2	87.1
Investm. amount (in k EUR)	79.2	40.8	39.3	39.7	39.3	39.2	46.1	70.7
Income (in k EUR)	25.5	27.9	35.1	53.3	25.5	30.3	38.1	45.2
Housing equity (in k EUR)	1.7	5.6	10.7	26.3	7.7	7.4	7.6	6.1
Housing debt (in k EUR)	0.0	0.0	0.0	0.0	0.1	0.1	0.0	0.0
Food expens. p.a. (in k EUR)	8.5	9.0	10.0	12.2	8.3	9.2	10.5	11.6
Share of secure inc.	0.8	0.9	0.9	0.9	0.9	0.9	0.9	0.9
Risk pref. stated {0,1}	1.0	0.8	0.7	0.7	0.8	0.8	0.8	0.9
Risk tolerance [0,5]	1.9	2.2	2.5	3.0	1.5	1.9	2.4	3.9
Trading exp.<1yr {0,1}	0.4	0.4	0.3	0.3	0.4	0.4	0.3	0.3
Trading exp. >3yr {0,1}	0.3	0.3	0.3	0.4	0.3	0.3	0.4	0.3
Age	50.5	48.8	44.3	38.3	59.8	47.0	40.4	34.6
Female {0,1}	0.5	0.5	0.5	0.5	0.5	0.5	0.5	0.5
Married {0,1}	0.7	0.7	0.8	0.7	0.8	0.8	0.7	0.7
Kids [0,4]	1.5	1.5	1.7	1.5	1.1	1.6	1.7	1.8
Certified professional {0,1}	0.2	0.3	0.2	0.2	0.2	0.2	0.3	0.3
High school degree {0,1}	0.3	0.3	0.2	0.2	0.3	0.3	0.3	0.2
College degree {0,1}	0.2	0.2	0.3	0.3	0.2	0.3	0.3	0.2
Self employed {0,1}	0.1	0.2	0.1	0.1	0.2	0.1	0.1	0.0
Retired {0,1}	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1

Table 8.
Household profiles by quartiles of the recommendation heterogeneity (standard deviation)
Household profiles are sorted into quartiles of the recommendation heterogeneity (standard deviation) across lay and professional advisors, represented by columns. The first row contains the quartile means of recommendation standard deviations, the following rows report household profile characteristics averaged within the corresponding quartile. For example, the quartile of households with the highest recommendation heterogeneity among lay advisors is shown in column (4).

	Lay Advisors				Professional Advisors			
	Q1	Q2	Q3	Q4	Q1	Q2	Q3	Q4
Risky Share Std. Deviation	24.2	28.7	31.2	34.7	21.0	28.6	31.7	34.9
Investm. amount (in k EUR)	90.6	51.6	34.7	31.0	82.2	53.3	35.1	32.7
Income (in k EUR)	48.9	38.1	31.9	22.4	40.7	34.0	32.6	29.4
Housing equity (in k EUR)	8.1	7.4	6.3	7.1	7.8	6.1	8.2	6.8
Housing debt (in k EUR)	0.0	0.0	0.0	0.0	0.0	0.1	0.0	0.0
Food expens. p.a. (in k EUR)	11.6	10.6	9.7	7.8	10.8	10.0	9.6	8.9
Share of secure inc.	0.9	0.8	0.8	0.9	0.8	0.8	0.9	0.9
Risk pref. stated {0,1}	0.9	0.8	0.7	0.7	1.0	0.9	0.7	0.6
Risk tolerance [0,5]	3.2	2.6	2.0	1.8	3.6	2.6	1.8	1.6
Trading exp.<1yr {0,1}	0.3	0.3	0.3	0.3	0.4	0.4	0.3	0.3
Trading exp. >3yr {0,1}	0.3	0.3	0.3	0.3	0.3	0.3	0.3	0.4
Age	36.4	39.8	45.6	59.9	39.0	45.2	44.1	53.5
Female {0,1}	0.5	0.5	0.5	0.5	0.5	0.5	0.5	0.5
Married {0,1}	0.7	0.7	0.8	0.7	0.8	0.8	0.8	0.7
Kids [0,4]	1.7	1.8	1.8	0.9	1.9	1.8	1.7	0.9
Certified professional {0,1}	0.2	0.3	0.2	0.3	0.2	0.3	0.2	0.2
High school degree {0,1}	0.3	0.3	0.3	0.2	0.3	0.3	0.2	0.3
College degree {0,1}	0.2	0.2	0.3	0.3	0.2	0.2	0.3	0.3
Self employed {0,1}	0.0	0.0	0.0	0.4	0.0	0.0	0.1	0.4
Retired {0,1}	0.1	0.1	0.1	0.2	0.1	0.1	0.1	0.2

Table 9.
Regression t-test for differences in recommendation average and standard deviation between lay and professional advisors
The table reports OLS regression coefficients and corresponding p-values. Dependent variables are the average and standard deviations of predicted recommendations for risky asset shares across advisors for each of 4355 household profiles, separately for the two advisor groups. Column pairs (1) and (3) report results for the regression on a constant, column pairs (2) and (4) for the regressions on a constant and a dummy equal to one for the professional advisor group.

	Risky Share Average				Standard deviation			
	(1)		(2)		(3)		(4)	
	Coeff.	p-val	Coeff.	p-val	Coeff.	p-val	Coeff.	p-val
const	0.657	0.000	0.668	0.000	0.294	0.000	0.297	0.000
Prof. Advisor			−0.023	0.000			−0.007	0.000
R2		0.000		0.006		0.000		0.005
N		8710		8710		8710		8710

Table 10.
Househould profiles by advisor groups' high vs. low recommendation average
Household profiles are sorted by averages in the recommended risky asset share across advisors for each of the two advisor groups. We report the most extreme household profiles that receive high or low average recommendations from both advisor groups or a low average from one and a high average from the other group. Column (1) for example shows the household profile with the highest average in recommendations received from both professional advisors and lay advisors.

Prof./Lay	(1) high/high	(2) high/low	(3) low/high	(4) low/low
Prof Mean	97.9	87.4	34.4	22.3
Lay Mean	95.9	47.7	68.5	30.6
Investm. amount (in k EUR)	20.0	900.0	20.0	900.0
Income (in k EUR)	100.0	50.0	40.0	10.0
Housing equity (in k EUR)	700.0	125.0	500.0	200.0
Housing debt (in k EUR)	0.0	0.0	150.0	5.0
Food expens. p.a. (in k EUR)	32.5	17.0	10.5	5.0
Share of secure inc.	1.0	1.0	1.0	1.0
Risk pref. stated {0,1}	1.0	1.0	0.0	1.0
Risk tolerance [0,5]	5.0	2.0	0.0	1.0
Trading exp.<1yr {0,1}	0.0	1.0	0.0	0.0
Trading exp. >3yr {0,1}	1.0	0.0	0.0	0.0
Age	20.0	35.0	62.0	64.0
Female {0,1}	0.0	1.0	0.0	1.0
Married {0,1}	1.0	0.0	1.0	1.0
Kids [0,4]	0.0	3.0	2.0	1.0
Certified professional {0,1}	0.0	0.0	0.0	1.0
High school degree {0,1}	0.0	1.0	0.0	0.0
College degree {0,1}	1.0	0.0	0.0	0.0
Self employed {0,1}	0.0	0.0	0.0	0.0
Retired {0,1}	0.0	0.0	0.0	1.0

Table 11.
Househould profiles by advisor groups' high vs. low recommendation standard deviation

The table follows the same logic as table 10; but, instead of sorting by recommendation average, its sorted by the recommended risk asset shares' heterogeneity (standard deviation). Column (1) for example shows the household profile with the highest standard deviation in recommendations for professional advisors and lay advisors.

Prof./Lay	(1) high/high	(2) high/low	(3) low/high	(4) low/low
Prof SD	39.1	30.1	11.6	9.3
Lay SD	37.7	21.8	27.2	12.8
Investm. amount (in k EUR)	20.0	900.0	800.0	20.0
Income (in k EUR)	5.0	15.0	50.0	100.0
Housing equity (in k EUR)	350.0	0.0	0.0	700.0
Housing debt (in k EUR)	0.0	0.0	0.0	0.0
Food expens. p.a. (in k EUR)	5.5	1.5	9.0	32.5
Share of secure inc.	1.0	1.0	0.5	1.0
Risk pref. stated {0,1}	1.0	1.0	1.0	1.0
Risk tolerance [0,5]	5.0	1.0	5.0	5.0
Trading exp.<1yr {0,1}	0.0	0.0	0.0	0.0
Trading exp. >3yr {0,1}	0.0	0.0	1.0	1.0
Age	75.0	33.0	38.0	20.0
Female {0,1}	0.0	1.0	0.0	0.0
Married {0,1}	1.0	0.0	0.0	1.0
Kids [0,4]	0.0	0.0	4.0	0.0
Certified professional {0,1}	0.0	1.0	0.0	0.0
High school degree {0,1}	1.0	0.0	0.0	0.0
College degree {0,1}	0.0	0.0	0.0	1.0
Self employed {0,1}	1.0	0.0	0.0	0.0
Retired {0,1}	0.0	1.0	0.0	0.0

5. Conclusion

What patients and clients of financial advisors have in common is that they should receive prescriptions based on their current circumstances, tailored to their individual needs. Previous findings suggest that financial advisors recommend portfolios that are highly similar across their clients and based on their own private allocations (see Foerster et al. (2017)). Our large portfolio experiment reveals that advisors do adjust recommendations based on the demographic information provided on virtual clients. We find risk tolerance, age, and income to be the most important factors in the advisors' portfolio rules. Surprisingly, wealth, whether financial or real estate, only plays a secondary role when judging by the effect size.

We identify advisor characteristics that cause the largest variation in the portfolio allocation rules, i.e. age, return expectations and risk tolerance; these directly increase the recommended risky share by roughly 6% to 9% for an average client after an increase close to one standard deviation in the corresponding advisor characteristic. When incorporating the information on household demographics, advisors seem to disagree primarily on how to map financial wealth into risky asset shares.

We find larger heterogeneity in the predicted recommendations of professional advisors. This suggests that professional advisors adapt the optimal risk allocation more actively in response to changes in household characteristics than lay advisors do. Furthermore, there is no convincing evidence that professional advisors incorporate their beliefs or own investment strategies stronger into recommendations than lay advisors. This is an interesting finding that suggests that projecting one's own preferences on to others is a human trait not specific to the financial advisor profession. By contrast, Leuermann and Roth (2012) show that both professional and lay advisors use their own risk preferences as a reference point when predicting other people's risk preferences, but they find a stronger effect for non-professional advisors compared with young professional advisors. We predict recommendations for all advisor and household combinations, which reveals that lay advisors tend to recommend risky asset allocations more homogeneously in the sense that they are closer to the overall mean of recommendations across all household profiles in comparison with professional recommendations. This is reflected in the overall standard deviation, which is larger for the recommendations of professional advisors and is consistent with the argument that professional advisors put more effort into matching households and their characteristics with an optimal portfolio allocation. While earlier research emphasizes the "one-size-fits-all" heuristics of financial advisors (see Foerster et al. (2017)), in the setting of our experiment, professional advisors are better matchmakers than people from the street. This suggests that professional advice potentially represents value added for private investors.

A surprising finding is that households in the quartile with the lowest risky asset share recommendations of lay advisors show a large average in financial wealth of €80,000 compared with around €40,000 in the other quartiles. Rich investors therefore receive lower risky share recommendations from lay advisors and the effect is reversed for professional advisors. This raises the question if risk aversion decreases with wealth for professional advisors while it increases for regular people. Another possible explanation is that lay advisors have difficulties in abstracting from their own situation and, say, adjust the reference point for gains and losses to the situation observed for the household profile at hand. This would imply that regular people would benefit from financial advice especially when their own life circumstances are subject to considerable change. Investigating investors' ability to cope with such transition periods is certainly a topic for future research. In any case, the finding reflects the complicated role of financial wealth as outlined in section 3.

Finally, we find that there are household profiles, i.e. very young, risk tolerant, and rich investors, that unambiguously receive large risky asset share recommendations. On the other hand, poorer, more risk averse and older investors face a larger heterogeneity in recommendations. Unfortunately, those most in need of advice seem to face the largest uncertainty. This adds to the cost of participating in financial markets which keeps a substantial proportion of households from investing in stock markets, as discussed for example in Haliassos and Bertaut (1995). If clients are aware of the heterogeneity in recommendations, we also deliver a possible explanation why those most in need of advice are more likely to refuse it (see Bhattacharya et al. (2012)).

Bibliography

Ameriks, John, and Stephen P. Zeldes. 2004. “How Do Household Portfolio Shares Vary with Age?” *Working Paper.*

Barsky, R. B., F. T. Juster, M. S. Kimball, and M. D. Shapiro. 1997. “Preference Parameters and Behavioral Heterogeneity: An Experimental Approach in the Health and Retirement Study”. *Quarterly Journal of Economics* 112 (2): 537–579.

Bertaut, Carol C., and Martha Starr-McCluer. 2000. “Household Portfolios in the United States”. *SSRN Electronic Journal.*

Bhattacharya, Utpal, Andreas Hackethal, Simon Kaesler, Benjamin Loos, and Steffen Meyer. 2012. “Is Unbiased Financial Advice to Retail Investors Sufficient? Answers from a Large Field Study”. *Review of Financial Studies* 25 (4): 975–1032.

Bluethgen, Ralph, Andreas Gintschel, Andreas Hackethal, and Armin Mueller. 2008. “Financial Advice and Individual Investors’ Portfolios”. *Working Paper.*

Bodie, Zvi, and Dwight B. Crane. 1997. “Personal Investing: Advice, Theory, and Evidence from a Survey of TIAA-CREF Participants”. *Working Paper.*

Bodie, Zvi, Robert C. Merton, and William F. Samuelson. 1992. “Labor supply flexibility and portfolio choice in a life cycle model”. *Journal of Economic Dynamics and Control* 16 (3-4): 427–449.

Brunnermeier, Markus K., and Stefan Nagel. 2008. “Do Wealth Fluctuations Generate Time-Varying Risk Aversion? Micro-evidence on Individuals”. *American Economic Review* 98 (3): 713–736.

Calvet, Laurent E., and Paolo Sodini. 2014. “Twin Picks: Disentangling the Determinants of Risk-Taking in Household Portfolios”. *The Journal of Finance* 69 (2): 867–906.

Chetty, R. A.J., László Sándor, and Adam Szeidl. 2017. “The Effect of Housing on Portfolio Choice”. *The Journal of Finance* 72 (3): 1171–1212.

Chetty, Raj, and Adam Szeidl. 2012. *The Effect of Housing on Portfolio Choice: NBER Working Papers.* National Bureau of Economic Research, Inc.

Chiappori, Pierre-AndrÃ©, and Monica Paiella. 2011. “Relative Risk Aversion is Constant: Evidence From Panel Data”. *Journal of the European Economic Association* 9 (6): 1021–1052.

Christiansen, Charlotte, Juanna Schröter Joensen, and Jesper Rangvid. 2008. "Are Economists More Likely to Hold Stocks?" *Review of Finance* 12 (3): 465–496.

Cocco, João F. 2005. "Portfolio Choice in the Presence of Housing". *The Review of Financial Studies* 18 (2): 535–567.

Cocco, João F., Francisco J. Gomes, and Pascal J. Maenhout. 2005. "Consumption and Portfolio Choice over the Life Cycle". *Review of Financial Studies* 18 (2): 491–533.

Dalal, Dev K., and Michael J. Zickar. 2012. "Some Common Myths About Centering Predictor Variables in Moderated Multiple Regression and Polynomial Regression". *Organizational Research Methods* 15 (3): 339–362.

Deutsches Aktieninstitut. 2017. "Aktionärszahlen: Industry Study". Visited on September 18, 2018. `https://www.dai.de/files/dai_usercontent/dokumente/studien/2018-02-19%20Aktieninstitut%20Aktionaerszahlen%202017%20Web.pdf`.

Foerster, S., J. T. Linnainmaa, B. T. Melzer, and A. Previtero. 2017. "Retail Financial Advice: Does One Size Fit All?" *The Journal of Finance* 72 (4): 1441–1482.

Friend, Irwin, and Marshall E. Blume. 1975. "The Demand for Risky Assets". *American Economic Review* 65 (5): 900–922.

Frühwirth-Schnatter, Sylvia, and Regina Tüchler. 2008. "Bayesian parsimonious covariance estimation for hierarchical linear mixed models". *Statistics and Computing* 18 (1): 1–13.

Gomes, Francisco, and Alexander Michaelides. 2003. "Portfolio Choice With Internal Habit Formation: A Life-Cycle Model With Uninsurable Labor Income Risk". *Review of Economic Dynamics* 6 (4): 729–766.

Guiso, Luigi, and Monica Paiella. 2008. "Risk Aversion, Wealth, and Background Risk". *Journal of the European Economic Association, MIT Press* 6 (6): 1109–1150.

Guiso, Luigi, and Paolo Sodini. 2013. "Household Finance: An Emerging Field", 2:1397–1532. Handbook of the Economics of Finance. Elsevier.

Hackethal, Andreas, Michael Haliassos, and Tullio Jappelli. 2012. "Financial advisors: A case of babysitters?" *Journal of Banking & Finance* 36 (2): 509–524.

Haliassos, Michael, and Carol C. Bertaut. 1995. "Why Do So Few Hold Stocks?" *Economic Journal* 105 (432): 1110–1129.

Heaton, John, and Deborah Lucas. 2000a. "Portfolio Choice and Asset Prices: The Importance of Entrepreneurial Risk". *The Journal of Finance* 55 (3): 1163–1198.

———. 2000b. "Portfolio Choice in the Presence of Background Risk". *Economic Journal* 110 (460): 1–26.

Huang, Chi-fu, and Robert H. Litzenberger. 1988. *Foundations for financial economics.* North-Holland.

Inderst, Roman, and Marco Ottaviani. 2009. “Misselling through Agents”. *American Economic Review* 99 (3): 883–908.

King, Mervyn A., and Jonathan I. Leape. 1998. “Wealth and portfolio composition: Theory and evidence”. *Journal of Public Economics* 69 (2): 155–193.

Leuermann, Andrea, and Benjamin Roth. 2012. *Does Good Advice Come Cheap? On the Assessment of Risk Preferences in the Lab and in the Field: SOEPpapers.* DIW Berlin, The German Socio-Economic Panel (SOEP).

Linnainmaa, Juhani, Brian Melzer, and Alessandro Previtero. 2018. “The Misguided Beliefs of Financial Advisors”. *Journal of Finance (forthcoming).*

Mankiw, N.Gregory, and Stephen P. Zeldes. 1991. “The consumption of stockholders and non-stockholders”. *Journal of Financial Economics* 29 (1): 97–112.

Merton, Robert C. 1969. “Lifetime Portfolio Selection under Uncertainty: The Continuous-Time Case”. *The Review of Economics and Statistics* 51 (3): p 247–257.

Mossin, Jan. 1968. “Optimal multiperiod portfolio policies”. *Journal of Business* 41:215–229.

Mullainathan, Sendhil, Markus Noeth, and Antoinette Schoar. 2012. “The Market for Financial Advice: An Audit Study”. *NBER Working Paper,* number 17929.

Net Fonds and Toluna. 2015. “Servicemonitor Finanzen: Industry Study”. Visited on September 18, 2018. `https://service.netfonds.de/marketing/pdf/Studie_Servicemonitor_B2C_2015.pdf`.

Polkovnichenko, Valery. 2007. “Life-Cycle Portfolio Choice with Additive Habit Formation Preferences and Uninsurable Labor Income Risk”. *The Review of Financial Studies* 20 (1): p 83–124.

Postbank and Kantar Emnid. 2017. “Sparen 2017: Tops und Flops: Industry Study”. Visited on September 18, 2018. `https://www.postbank.de/postbank/pr_presseinformation_2017_10_30_sparen_2017_tops_und_flops.html`.

Rossi, Peter E., Greg M. Allenby, and Robert Edward McCulloch. 2005. *Bayesian statistics and marketing.* J. Wiley & Sons.

Samuelson, Paul A. 1969. “Lifetime Portfolio Selection by Dynamic Stochastic Programming”. *The Review of Economics and Statistics* 51 (3): 239–246.

Union Investment and Forsa. 2018. "Anlegerbarometer Q2". Visited on September 18, 2018. `https://unternehmen.union-investment.de/startseite-unternehmen/presseservice/pressemitteilungen/alle-pressemitteilungen/2018/Studie--Anlegerbarometer---zweites-Quartal-2018.html`.

Vissing-Jorgensen, Annette. 2002. "Limited Asset Market Participation and the Elasticity of Intertemporal Substitution". *Journal of Political Economy* 110 (4): 825–853.

Appendix A - MCMC sampler and estimation

A.I. The MCMC Sampler

MCMC Sampler

We use the same sampling algorithm as described in Frühwirth-Schnatter and Tüchler (2008), and add a selection step before the posterior draw of the vector $(\alpha\ \beta)$.

Priors

- elements in random effects indicator matrix: $\gamma^r \sim Beta(q_r + 1, d_c - q_r + 1)$
- Cholesky factor (selected elements, stacked): $C^\gamma \sim \mathcal{N}(a_0, \sigma_\varepsilon^2 A_0)$
- elements fixed effects indicator vector: $\gamma^f \sim Beta(q_f + 1, d_f - q_f + 1)$
- fixed effects and random effects' mean: $(\alpha\ \beta) \sim \mathcal{N}(b_0, B_0)$
- observation error variance: $\sigma_\varepsilon^2 \sim InvGamma(s_0/2, S_0/2)$

We will keep the priors on the random effect means and fixed effects, i.e. B_0 for $(\boldsymbol{\alpha}\ \boldsymbol{\beta})$ as well as for the random effects' covariance matrix, A_0 for $\boldsymbol{C}$, fairly uninformative, reporting results for different scalar model tuning parameters for the prior dispersion $p_{A_0} \geq 1000$ and $p_{B_0} \geq 1000$. The off-diagonal elements in A_0 and B_0 are kept at 0. As seen below p_{A_0} and p_{B_0} scale the diagonal elements of A_0 and B_0. For example, $p_{B_0} = 10,000$ would set the prior standard deviation of (α, β) to 100, which is large given that the dependent variable is on the log scale. The full set of prior parameters is:

$$
\begin{aligned}
a_0 &= \mathbf{0} \in \mathbb{R}^{q_r} \\
A_0 &= p_{A_0} \cdot \boldsymbol{I} \in \mathbb{R}^{q_r \times q_r} \\
b_0 &= \mathbf{0} \in \mathbb{R}^{d+q_f} \\
B_0 &= p_{B_0} \cdot \boldsymbol{I} \in \mathbb{R}^{(d_r+q_f)\times(d+q_f)} \\
s_0 &= 0.1 \\
S_0 &= 0.1
\end{aligned}
$$

MCMC scheme

1. sample $\boldsymbol{\gamma}^f_m | \boldsymbol{\gamma}^r_{\setminus m}, \boldsymbol{\gamma}^r, \boldsymbol{C}^\gamma, \sigma_\varepsilon^2, \boldsymbol{y}$ from a discrete density with two realizations ($\boldsymbol{\gamma}^f_m$ denotes the

mth element in $\boldsymbol{\gamma}^f$, $\boldsymbol{\gamma}^f_{\backslash m}$ all indicators excluding the mth element, see Frühwirth-Schnatter and Tüchler (2008) for details and below for the marginal likelihood used in this step).

2. sample $(\alpha^\gamma\ \beta)|\boldsymbol{\gamma}^f, \boldsymbol{\gamma}^r, \boldsymbol{C}^\gamma, \sigma^2_\varepsilon, \boldsymbol{y} \sim \mathcal{N}_{q_f+d_r}(\boldsymbol{B}_N\boldsymbol{b}_N, \boldsymbol{B}_N)$, $p(\boldsymbol{\alpha}, \boldsymbol{\beta}|\boldsymbol{\gamma}, C^\gamma, \sigma^2_\varepsilon, \boldsymbol{y})$
   with $b_N = \sum_{a=1}^N [\boldsymbol{X}^f_a \boldsymbol{X}^r_a]'(\boldsymbol{X}^r_a \boldsymbol{Q}(\boldsymbol{X}^r_a)' + \sigma^2_\varepsilon \boldsymbol{I}_5)^{-1}\boldsymbol{y}_a + \boldsymbol{B}_0^{-1}\boldsymbol{b}_0$,
   $B_N^{-1} = \sum_{a=1}^N [\boldsymbol{X}^f_a \boldsymbol{X}^r_a]'(\boldsymbol{X}^r_a \boldsymbol{Q}(\boldsymbol{X}^r_a)' + \sigma^2_\varepsilon \boldsymbol{I}_5)^{-1}[\boldsymbol{X}^f_a \boldsymbol{X}^r_a] + \boldsymbol{B}_0^{-1}$
3. sample $z|\boldsymbol{\gamma}^f, \boldsymbol{\alpha}^\gamma, \boldsymbol{\gamma}^r, C^\gamma, \boldsymbol{\beta}, \sigma^2_\varepsilon, \boldsymbol{y} \sim \mathcal{N}_{d_r}(\boldsymbol{P}_a\boldsymbol{p}_a, \boldsymbol{P}_a)$
   with $\boldsymbol{p}_a = \sigma_\varepsilon^{-2}(\boldsymbol{X}^r_a C)'(\boldsymbol{y}_a - \boldsymbol{X}^f_a\boldsymbol{\alpha} - \boldsymbol{X}^r_a\boldsymbol{\beta})$,
   $\boldsymbol{P}_a^{-1} = \sigma_\varepsilon^{-2}(\boldsymbol{X}^r_a C)'(\boldsymbol{X}^r_a C) + \boldsymbol{I}_{d_r}$
4. sample $\boldsymbol{\gamma}^r_{lm}|\boldsymbol{\gamma}^r_{\backslash lm}, \boldsymbol{\gamma}^f, \alpha^\gamma, \boldsymbol{\beta}, \boldsymbol{z}, \sigma^2_\varepsilon, \boldsymbol{y}$ from a discrete density with two realizations ($\boldsymbol{\gamma}^r_{lm}$ denotes the element in row l and column m of $\boldsymbol{\gamma}^r$, $\boldsymbol{\gamma}^r_{\backslash lm}$ all indicators excluding the lmth element, see Frühwirth-Schnatter and Tüchler (2008) for details).
5. sample $\boldsymbol{C}^\gamma|\boldsymbol{\gamma}^f, \boldsymbol{\alpha}^\gamma, \boldsymbol{\gamma}^r, \boldsymbol{\beta}, \boldsymbol{z}, \sigma^2_\varepsilon, \boldsymbol{y} \sim N_{q_r}(\boldsymbol{a}_N, \sigma^2_\varepsilon \boldsymbol{A}_N)$
   with $(\boldsymbol{A}_N^{\text{ran}})^{-1} = \sum_{a=1}^N (\boldsymbol{W}^r_a)'\boldsymbol{W}^r_a$, $\ b = 1/\sum_{a=1}^N 5$,
   $\boldsymbol{a}_N = \boldsymbol{A}_N \left(\sum_{a=1}^N (\boldsymbol{W}^r_a)'(\boldsymbol{y}_a - \boldsymbol{X}^f_a\boldsymbol{\alpha} - \boldsymbol{X}^r_a\boldsymbol{\beta})\right)$
6. sample $\sigma^2_\varepsilon|\boldsymbol{\gamma}^f, \alpha^\gamma, \boldsymbol{\gamma}^r C^\gamma, \beta, \boldsymbol{z}, \boldsymbol{y} \sim \mathcal{G}^{-1}(s_N/2, S_N/2)$
   with $s_N = \sum_{a=1}^N T_a + s_0$,
   $S_N = S_0 + S^\gamma + (\boldsymbol{a}_N - \boldsymbol{a}_0)'\boldsymbol{A}_0^{-1}(\boldsymbol{a}_N - \boldsymbol{a}_0)$
   $S^\gamma = \sum_{a=1}^N \|\boldsymbol{y}_a - \boldsymbol{W}^r_a\boldsymbol{a}_N - \boldsymbol{X}^f_a\boldsymbol{\alpha} - \boldsymbol{X}^r_a\boldsymbol{\beta}\|_2$

Marginal likelihood for fixed effects selection step

In order to sample the fixed effects indicator, we derive the likelihood function marginal with respect to $(\boldsymbol{\alpha}\ \boldsymbol{\beta})$ and the random effects $\boldsymbol{z}$, since the non-normalized posterior of $(\boldsymbol{\alpha}\ \boldsymbol{\beta})$ integrated over $(\boldsymbol{\alpha}\ \boldsymbol{\beta})$ is equal to the marginal likelihood.
Define:

$$\boldsymbol{\alpha}^\star = [\boldsymbol{\alpha}', \boldsymbol{\beta}']' \tag{A.1}$$

$$\boldsymbol{D}'\boldsymbol{D} = (\boldsymbol{X}^r_a \boldsymbol{Q}(\boldsymbol{X}^r_a)' + \sigma^2_\varepsilon \boldsymbol{I}_5)^{-1} \tag{A.2}$$

$$\boldsymbol{y}^\star_a = (\boldsymbol{y}'_a\boldsymbol{D}')' \tag{A.3}$$

$$[\boldsymbol{X}^f_a \boldsymbol{X}^r_a]^\star = ([\boldsymbol{X}^f_a \boldsymbol{X}^r_a]'\boldsymbol{D}')' \tag{A.4}$$

Such that

$$\begin{aligned}&(\boldsymbol{y}^\star_a - [\boldsymbol{X}^f_a \boldsymbol{X}^r_a]^\star[\boldsymbol{\alpha}', \boldsymbol{\beta}']')'(\boldsymbol{y}^\star_a - [\boldsymbol{X}^f_a \boldsymbol{X}^r_a]^\star[\boldsymbol{\alpha}', \boldsymbol{\beta}']') = \\ &(\boldsymbol{y}_a - [\boldsymbol{X}^f_a \boldsymbol{X}^r_a][\boldsymbol{\alpha}', \boldsymbol{\beta}']')'(\boldsymbol{X}^r_a \boldsymbol{Q}(\boldsymbol{X}^r_a)' + \sigma^2_\varepsilon \boldsymbol{I}_{T_a})^{-1}(\boldsymbol{y}_a - [\boldsymbol{X}^f_a \boldsymbol{X}^r_a][\boldsymbol{\alpha}', \boldsymbol{\beta}']')\end{aligned} \tag{A.5}$$

With

$$\boldsymbol{y}^\star = [(\boldsymbol{y}^\star_1)', \ldots, (\boldsymbol{y}^\star_a)', \ldots, (\boldsymbol{y}^\star_N)']' \tag{A.6}$$

$$\boldsymbol{X}^\star = [([\boldsymbol{X}^f_1 \boldsymbol{X}^r_1]^\star)', \ldots, ([\boldsymbol{X}^f_a \boldsymbol{X}^r_a]^\star)'), \ldots, ([\boldsymbol{X}^f_N \boldsymbol{X}^r_N]^\star)']' \tag{A.7}$$

multiplying likelihood time prior

$$\begin{aligned} p(\boldsymbol{\alpha}^{\star}|\boldsymbol{y}^{\star},\ldots) &\propto p(\boldsymbol{y}^{\star}|\boldsymbol{\alpha}^{\star}...) \times p(\boldsymbol{\alpha}^{\star}) \\ &\propto exp\left[\frac{1}{2}\left(\boldsymbol{y}^{\star}-\boldsymbol{X}^{\star}\boldsymbol{\alpha}^{\star}\right)'\left(\boldsymbol{y}^{\star}-\boldsymbol{X}^{\star}\boldsymbol{\alpha}^{\star}\right)\right] \\ &\times exp\left[\frac{1}{2}\left(\boldsymbol{\alpha}^{\star}-\boldsymbol{b}^{0}\right)'\boldsymbol{B}^{0}\left(\boldsymbol{\alpha}^{\star}-\boldsymbol{b}^{0}\right)\right] \end{aligned} \tag{A.8}$$

and integration over $\boldsymbol{\alpha}^{\star}$ yields marginal likelihood:

$$\begin{aligned} &p(\boldsymbol{y}^{\star}|\boldsymbol{A}^{0},\boldsymbol{a}^{0},\ldots) = \\ &\quad |\boldsymbol{A}^{0}|^{.5}\left|\left(\tilde{\boldsymbol{B}}\right)\right|^{(-.5)} exp\left(\frac{-n\tilde{s}^{2}}{2}\right) \end{aligned} \tag{A.9}$$

with

$$n\tilde{s}^{2} = \left(\boldsymbol{y}^{\star}-\boldsymbol{X}^{\star}\tilde{\boldsymbol{\alpha}}^{\star}\right)'\left(\boldsymbol{y}^{\star}-\boldsymbol{X}^{\star}\tilde{\boldsymbol{\alpha}}^{\star}\right)+\left(\tilde{\boldsymbol{\alpha}}^{\star}-\boldsymbol{b}^{0}\right)'\boldsymbol{B}^{0}\left(\tilde{\boldsymbol{\alpha}}^{\star}-\boldsymbol{b}^{0}\right) \tag{A.10}$$

$$\tilde{\boldsymbol{B}} = \left((\boldsymbol{X}^{\star})'\boldsymbol{X}^{\star}+\boldsymbol{B}^{0}\right) \tag{A.11}$$

$$\tilde{\boldsymbol{\alpha}}^{\star} = \tilde{\boldsymbol{B}}^{-1}\left((\boldsymbol{X}^{\star})'\boldsymbol{y}^{\star}+\boldsymbol{B}^{0}\boldsymbol{b}^{0}\right) \tag{A.12}$$

$\boldsymbol{X}^{\star}$ contains only currently active columns as per $\{\gamma^{f}\}$. The dimensions of $\boldsymbol{a}^{0}$ and $\boldsymbol{A}^{0}$ are adjusted accordingly.

A.II. Sampler Tests

To test validity and performance of the estimation model and the MCMC, we first compare different sets of prior dispersion for $(\boldsymbol{\alpha}\ \boldsymbol{\beta})$ and $\boldsymbol{C}$ represented by the scaling parameter B_0 and A_0 respectively. To this end we generate new response data $\boldsymbol{y}_{a,h}$ by using the real data design matrices, see tables A.2 and A.3, and set coefficients to known values. Depending on the performance of different prior sets, we choose the two sets that seem most suitable for the data at hand and check the convergence and accuracy of the MCMC sampler by evaluating the time series plots for the parameter estimates.

We include $d_h = 20$ client profile variables, together with the constant, the random effects design matrix has $d_r = d_h + 1 = 21$ columns. With $d_a = 12$ variables on financial advisors' characteristics, we have 12 non-interaction and 240 interaction terms summing up to $d_f = 252$ columns for the fixed effects design matrix.

The elements in the parameter vector $(\alpha\ \beta)$ are equal to a repeating sequence of $(1.0, 1.5, 2.0, 2.5, 0.0)$. We restrict random effects on the first 5 random variables, setting all elements with row and column indices greater than 5 to zero. The upper left 5-dimensional submatrix of $\boldsymbol{C}$ receives non-zero elements in the lower triangular region, where off-diagonal elements are set to 1.5 and elements on the diagonal set to 2. The advisor specific fixed effects are randomly drawn from a normal distribution $\mathcal{N}_5(\mathbf{0}, \boldsymbol{I}_5)$ for each advisor a. Finally, σ_ε^2 is set to 9. The values are set arbitrarily, since our analyses serves only the purpose of gaining insights into the choice of prior dispersion, convergence and overall behavior of our Bayesian sampler given the real covariate data.

Table A.1.
Selection ratio of correctly excluded and correctly included (selected) regression coefficients

The table reports selection ratios by prior parameter set. Ratios are percentages of the total count of coefficients in the specific category (rows of this table) and over all simulation iterations. We ran the MCMC sampler using 100.000 iterations for each set of priors keeping every tenth iteration and dropping 50 percent as burn-in. This results in 5000 observations of each model parameter's posterior distribution.

Prior scalar p_{B_0} on $(\boldsymbol{\alpha}\ \boldsymbol{\beta})$	100	1,000	100	1,000	10,000	1,000	10,000
Prior scalar p_{A_0} on $(\boldsymbol{C})$	100	100	1,000	1,000	1,000	10,000	10,000
Selection on $\boldsymbol{\alpha}$: Correctly...							
excluded coefficients	1%	10%	1%	11%	37%	11%	37%
included coefficients	99%	97%	99%	97%	94%	97%	94%
Selection on $\boldsymbol{C}$: Correctly...							
excluded coefficients	98%	99%	97%	9%	98%	99%	99%
included coefficients	60%	60%	60%	60%	60%	60%	60%
Sim. iterations	5,000	5,000	5,000	5,000	5,000	5,000	5,000

In table A.1 we see the ratio of correctly excluding zero coefficients and correctly including non-zero coefficients. It seems evident that increasing prior dispersion for the fixed effects (B_0) leads to a better ratio of excluding variables that do not have an effect on the dependent variable (a coefficient that should be zero). Comparing for example the first and last column, with B_0 equal to 100 versus 10 000, the detection ratio of fixed effects that should be excluded increases from 1% to 37%, while the correct inclusion of non-zero fixed effects decreases only from 99% to 94%. As expected, varying B_0, the fixed effects dispersion, has no effect on random effects selection. Likewise, A_0 has no effect on fixed effects selection. In contrast to the fixed effects, varying the prior dispersion for the random effects does not seem to substantially influence the ratio of either excluding zero coefficients or including non-zero coefficients, the former stays above 90% while the latter stays at 60%. Given these results we preferably use the prior set $B_0 = 10\,000, A_0 = 10\,000$ for our estimation model.

A.III. Model inputs and convergence

In this section we illustrate that our MCMC chains indeed converge, even in a very large model with 263 coefficients in the alpha vector alone. Furthermore, we provide an overview of the variables used in the regression, and specifically their scaling, in table A.3 for household characteristics and in table A.2 for advisor characteristics. All predictors are centered by the population mean to reduce collinearity (see e.g. Dalal and Zickar (2012)) and make the coefficients on first order effects interpretable at the population average. We ran the MCMC sampler using four chains, each with 100.000 iterations, keeping every tenth iteration and dropping 75 percent as burn-in. This results in 10000 observations of each model parameter's posterior distribution.

Table A.2.
List of variables in the advisor fixed effects design matrix x_a
The table reports all avisor variables that we use, their abbreviations as well as the scaling for regressions. Appendix Appendix B section II provides the underlying survey questions.

j	Variable $x_{a,j}$	Regr. Abbrev.	Data type
1	Professional	Prof.Adv.	binary
2	Age	Age	integer [18,100]
3	Male	Male	binary
4	College	College	binary
5	Income	ln.Inc.	cont. log(x+1)
6	Real Estate Wealth	ln.RE.Wlth	cont. log(x+1)
7	Financial Wealth	ln.Fin.Wlth	cont. log(x+1)
8	Own Allocation	OwnAlloc.	cont. log(x+0.001)
9	Return Expectations 10yrs	Ret.Exp.10yrs	integer [0,100]
10	Return Expectations 1yr	Ret.Exp.1yr	integer [0,100]
11	Risk Tol. Bomb Game	RiskTol.B.	integer [0,100]
12	Patience	Patience	integer [0,5]
	Variable $y_{a,h}$	Data type	
	Risky Asset Recommendation	cont. log(x+0.001)	

In figures A.1 and A.2 we show the time series of the simulation iterations for coefficients in $\boldsymbol{\alpha}$, $\boldsymbol{\beta}$, and $\boldsymbol{Q}$. To observe the complete chain, we do not dispose iterations for burn-in. We only show every 100th iteration to make individual timelines visible. The MCMC chain seems to converge very fast. Coefficients in beta and Q converge after about 5000 iterations. Coefficients in alpha seem to reach their stable range equally fast. Given that there is always at least some probability that a coefficient in alpha is removed from the model and set to zero, it is not surprising that estimates in alpha seem to be less stable than in beta.

In contrast to the interaction model, we report selection rates on the first set of coefficients in

Table A.3.
List of variables in the virtual client profile design matrix x_h
The table reports all household variables that we use, their abbreviations as well as the scaling for regressions. Appendix Appendix B section I and II provide further details on all characteristics.

i	Variable $x_{h,i}$	Regr. Abbrev.	Data type
1	Investment Amount	ln.Inv.Amount	cont. log(x+1)
2	Income	ln.Income	cont. log(x+1)
3	Real Estate Wealth	ln.RE.Wlth	cont. log(x+1)
4	Real Estate Debt	ln.RE.Dbt	cont. log(x+1)
5	Food Expenditure	Food Expnd.	cont. log(x+1)
6	Safe Income	Safe Inc.	fraction [0,1]
7	Risk Tolerence provided	RiskTol.Prov.	binary
8	Risk Tolerance	RiskTol.	integer [0,5]
9	Inv. Experience < 1 year	Inv.Exp.<1yr	binary
10	Inv. Experience > 3 years	Inv.Exp.>3yrs	binary
11	Age	Age	integer [0,100]
12	Age Sqrd.	Age Sqrd.	integer [0,10000]
13	Male	Male	binary
14	Married	Married	binary
15	Kids	Kids	integer [1,100]
16	Professional Train.	Prof.Train.	binary
17	A Levels	ALevels	binary
18	College	College	binary
19	Retired	Retired	binary
20	Self Employed	SelfEmpl.	binary

the alpha vector. Panel A in table 6 includes estimates on advisor characteristics and their interaction with a dummy for professional advisors. The first column shows that important advisor characteristics such as age, own allocation, and return expectation are usually not excluded from the model with selection rates at or above 99%, meaning that coefficients were estimated to be different from zero in 99% of iterations during the MCMC simulations.

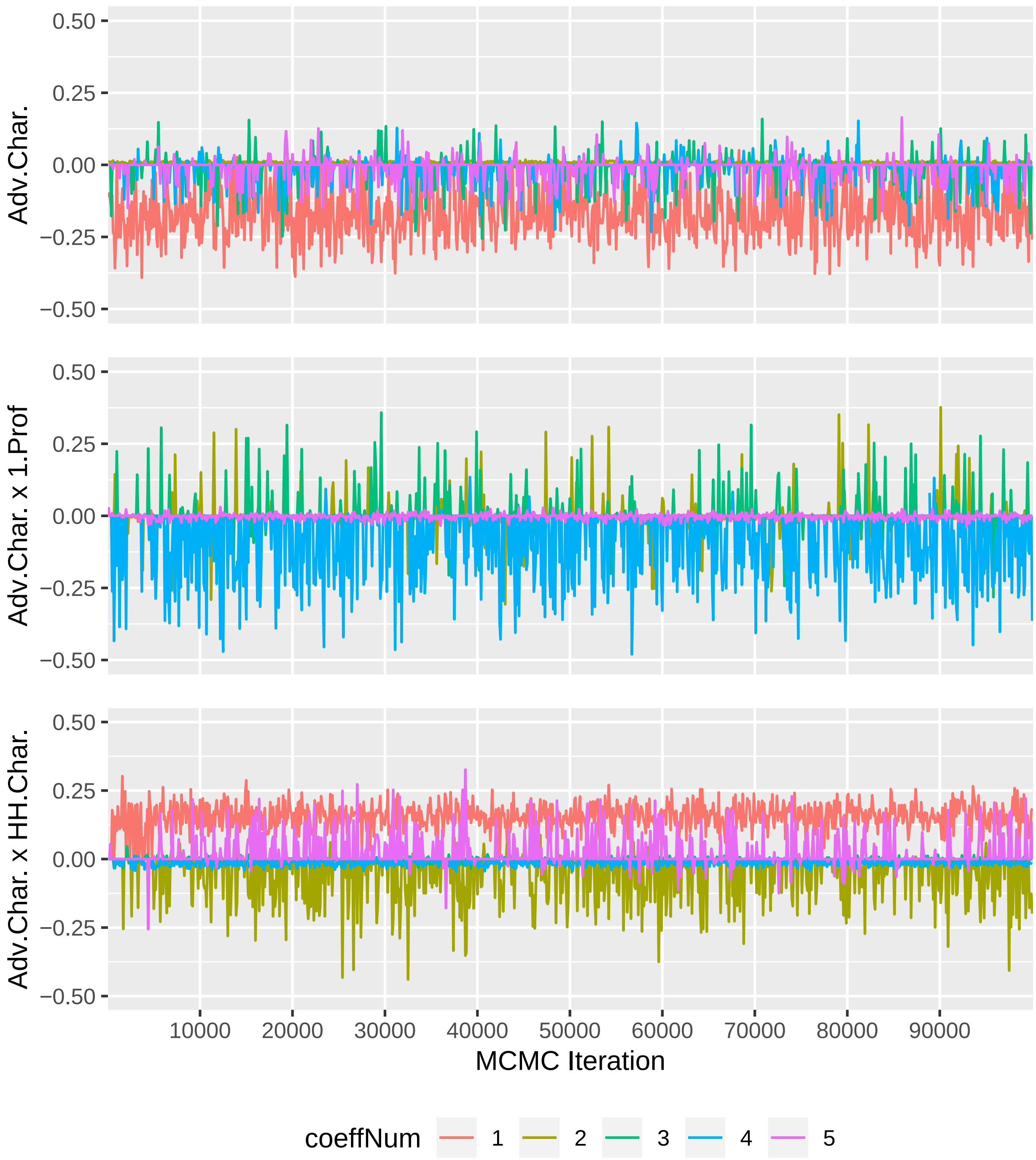


Figure A.1.
Alpha coefficient estimates over MCMC iterations.
The graphic displays the time series of coefficient estimates during simulation iterations for the first of four MCMC chains. We display every 100th of the 100,000 iterations and show only data points on the first five coefficients of the following categories: The upper panel includes coefficients on advisor characteristics, the panel in the middle includes coefficients on the interactions between advisor characteristics and the dummy for professional advisors, and the panel at the bottom shows coefficients on interactions between household and advisor characteristics.

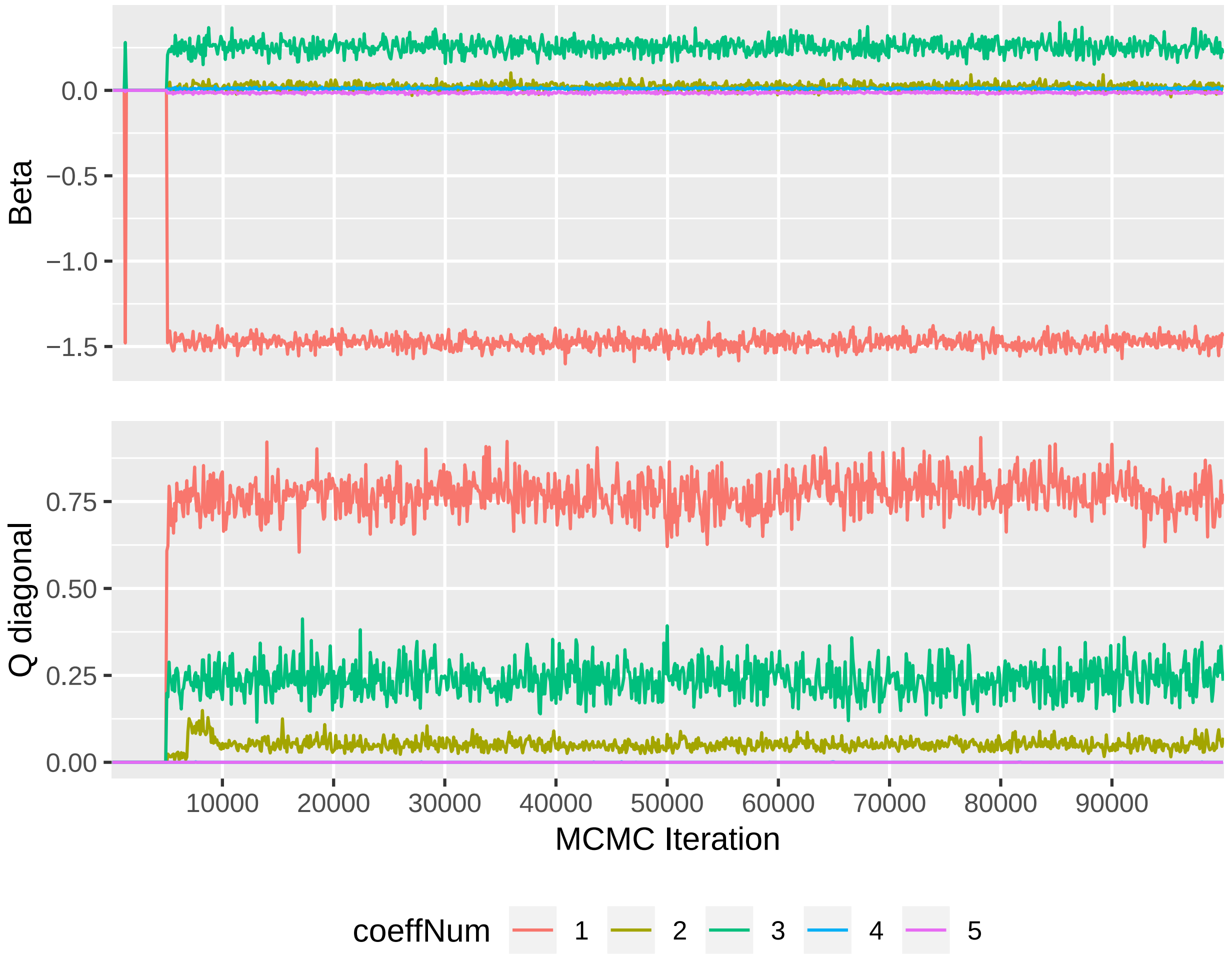


Figure A.2.
Beta and diagonal(Q) coefficient estimates over MCMC iterations.
The graphic displays the time series of coefficient estimates during simulation iterations for the first of four MCMC chains. We display every 100th of the 100,000 iterations and show only data points on the first five coefficients of the random effect means in the upper panel and diagonal elements of the Q-matrix in the lower panel.

Appendix B - Survey details

In this section of the appendix we first provide details on how investor profiles were generated. We abstained from generating completely random profiles to rule out implausible combinations such as huge amounts of real estate debt without real estate wealth. Furthermore, we provide two full transcripts for the survey based experiments on both lay and professional advisors that include notes on the rationale of choosing and incorporating the different questions in the survey.

B.I. Quasi-random generation of investor profiles

age	uniformly distributed min=20, max=75
sex	binomial p=0.5
married	binomial p=0.5
kids	uniformly distributed min=0 max=4
education	4 categories with equal probabilities (10yrs, 12yrs, professional diploma, college)
job	ages $<$ 55: self-employed with p=0.2; ages 55-64 self-empl. with p=0.2, retired with p=0.2; ages$>$ 64 retired with p=1
risk pref.	5 categories with equal probabilities
income	truncated normal: mean=40,000 Euro; SD=100,000; min=20,000; max=1,500,000
food spend.	5000 + c*income with $c \in (0.10, 0.25)$
real estate	uniformly distr. min=0, max=income*2
mortgages	uniformly distr. min = 0, max=(realestate wealth)*0.8
other debt	uniformly distr. min=0, max=$max\{income, 0.25^{*}inv.amount\}$
inv. exp.	3 categories with equal probabilities
inv. amount	uniformly distr. min=5000 max=income*20

B.II. Transcript Professional Advisors

1. Entrance Questionnaire - Characteristics of the Participants

1.1. Age

1. ______ [numerical input]

1.2. Years of experience in current profession

1. ______ [numerical input]

1.3. How are you registered in the (German) trade regulation act?

1. §34d Insurance broker
2. §34e Insurance consultant
3. §34f Financial investment broker
4. §34h Honorary financial consultant
5. No registration
6. Other: ______

1.4. Self-Employed

1. yes
2. no

1.5. Education

1. Highschool (10 years)
2. Highschool with higher education entrance qualification
3. Professional diploma ("Berufsausbildung")
4. Bachelor's degree
5. Master's degree
6. Doctoral degree
7. Other: ______
   [single choice]

1.6. Qualifications

1. Commercial professional diploma
2. Business administrator
3. Professional advisor
4. Chartered Financial Analyst (CFA)
5. Certified Financial Planner (CFP)
6. Applied Investment Management (AIM)
7. Certified International Investment Analyst (CIIA)
8. Other:______
   [single choice]

1.7. What percentage of your clients is paying you on a honorary basis?

1. ______

1.8. In your opinion, why are there so less registered honorary financial consultants?

1. ______

1.9. How can the acceptance of honorary based consulting be increased for suppliers and clients?

1. ______

1.10. How could statutory provisions be improved?

1. ______

1.11. How high would the honorarium be for one hour of consulting to work economically viable?

1. ______ Euro [numerical input]

1.12. How are you compensated on an average for your services as financial advisor and how high is your compensation?

1. fixed/value-proportional commission per sold product ______
2. performance fee ______
3. fixed fee per consultation/per hour ______
4. commission proportional to money invested by client ______
5. (recurring) commission proportional to profits gained ______
6. other ______

[multiple choice possible]

1.13. How high is the average investment of the typical client?

1. ______ Euro [numerical input]

1.14. In your experience, how much is a client willing to pay for an hour of advisory on a honorary basis?

1. from ______ Euro [numerical input]
2. to ______ Euro [numerical input]

1.15. What is the minimum investment amount that makes a customized consultation on a honorary basis is viable?

1. ______ Euro [numerical input]

2. Expected Returns

The second question group aims at collecting the return expectations for the risky and riskless asset. The risky asset is defined as an investment in a fund that mirrors the German DAX index. The riskless asset is a money market investment, with maturities up to a year.
Returns are assumed to be normally i.i.d., therefore it suffices to collect the subjective probabilities about returns exceeding two different thresholds.
To collect information about expected returns over different horizons and enable testing a market timing hypothesis, the first two questions will cover a return over one year and questions 3 and 4 will cover the average annual return over the next 10 years. Question 5 and 6 simply ask which return is expected for the riskless investment for one year and as an average over the next 10 years. The introduction text is loosely adapted from the HRS questionnaire, while the elicitation method for expected mean and standard deviation of the risky asset return is based on Kezdi and Willis (2009).

Description

We are interested in how you think about investments in the stock market. We have some questions about how much someone might make or lose from an investment in the stock market.
We want to know what you think the chances are of how much you might gain or lose on that

investment if you were to make it. Your answers can range from zero to one hundred, where zero means there is absolutely no chance, and excactly one hundred means that it is absolutely certain. (For example, when weather forecasters report the chance of rain, a number like 20 percent means "not much chance", a number between 45 and 55 percent means "a pretty even chance", and a number like 80 percent means "a very good chance.")
Suppose you left the € 10,000 in the mutual fund for one year, and didn't take out any dividends or interest, and then after one year you cashed in the mutual fund and took everything out. Assume that there are no commissions or fees for buying or selling this fund.

2.1. Return probability 1 year, more than 0 % return With which percentage probability do you expect to receive more than € 10,000 if you sell your mutual fund shares. I.e., you receive a net profit of more than € 0.

1. ______ % [numerical input]

restriction: 0<=answer2.1<=100

2.2. Return probability 1 year, more than 10 % return With which percentage probability do you expect the fund investment to have increased in value by more than 10%? I.e., if you sell your mutual fund shares you would receive more than € 11,000 - the net profit is more than € 1,000.

1. ______ % [numerical input]

restriction: 0<=answer2.2 < answer2.1<=100

2.3. Return probability 10 year annual average, more than 0 % return With which percentage probability do you expect that the mutual fund will average an annual return of more that 0% per year over the next 10 years - the net profit is positive.

1. ______ % [numerical input]

restriction: 0<=answer2.3<=100

2.4. Return probability 10 year annual average, more than 10 % return With which percentage probability do you expect that the mutual fund will average an annual return of more that 10% per year over the next 10 years. I.e. after 10 years of investment the net profit is more than € 15,937. $(10,000 * (1.10)^1 0 \approx 25,937)$

1. ______ % [numerical input]

restriction: 0<=answer2.4 < answer2.3<=100

2.5. Riskless asset return 1 year Please state the annual interest that you expect to earn on a riskless money market investment.

1. ______ % [numerical input]

2.6. Riskless asset return 10 years Please state the average annual interest that you expect to earn for the next 10 years on a riskless money market investment.

1. ______ % [numerical input]

2.7. Confidence On a scale from 1 to 5, how confident do you feel about the expectations you stated in the previous questions (with 1 meaning very unconfident and 5 meaning absolute confident).

1. ______ [numerical input]

3. Risk preferences

Risk preferences are measured with two questions. The first aiming at directly capturing the Arrow-Pratt measure of relative risk aversion. This is achieved by assigning the Bomb Risk Elicitation Task (BRET) as proposed by Crosetto and Filippin (2012). The second is a qualitative question that resembles the risk elicitation used by financial services providers in Germany to fulfill legal requirements (WPHG[1]) followed by a question about the participants patience.

3.1. BRET Imagine the following:

- There are 100 boxes. Inside of 99 of these boxes there is € 1.
- But inside of a single randomly selected box there is a bomb.
- You are allowed to open as many boxes as you want. Therefore you choose a number between 1 and 100.
- For each box, which you open that does not contain the bomb, you will receive a dollar. That means you would receive € 20, if you choose to open 20 boxes and non of them contains the bomb.
- If, unfortunately, the bomb is inside one of the boxes that you have chosen, you will lose everything and receive € 0.

How many boxes would you open?

1. ______ % [numerical input]

restriction: 1<=answer3.1<=100//

3.2. Qualitative risk elicitation question Please choose from below the statement that describes your own investment behavior best.

1. You put emphasis on save returns and do not want to take on risks associated with asset price fluctuations.
2. You put emphasis on stable returns and you are willing to take on only moderate risk associated with asset price fluctuations.
3. You put emphasis on returns above capital market interest rates and you are willing to take on certain risk of loss.
4. You put emphasis on increased returns especially through asset price gains and you are willing to take on considerable risk to of loss.

1. WPHG = "Wertpapier Handelsgesetz" *transl.: securities trading law*

5. You have high expectation for returns and you are willing to take on the corresponding high risk of loss.
   [single choice]

3.3. Patience On a scale from 0 to 10, how patient do you assess yourself? (With 0 meaning very impatient and 10 meaning absolute patient.)

1. ______ [numerical input]

4. Portfolio recommendations

The set-up for groups 4 to 8 is identical. Each group contains a single question. The participants will see a client profile, which represents a client that needs portfolio advice. The recommendation is solely in the form of a risky asset portfolio share. That is, the participants will have to allocate a certain Euro amount of the client's financial assets which is recommended to be invested in stocks (a vehicle mirroring the DAX index).

Description

Consider the following investor profile.
Imagine you were an investor in the situation described below. As only stocks and a money market investment are considered, you consequently need to know how much of your investment amount to allocate to stocks. Assume that the financial wealth of the whole household is being considered. Therefore the investment amount is the sum of all cash, savings, and deposit accounts as well as stocks and treasury bonds. Where applicable all other characteristics are household level aggregates as well. Do not mind if some of the situations seem unrealistic (like a 21 year old young entrepreneur with an inheritance of € 1,000,000 and an substantial amount of real estates). The investment amount does not include other wealth relevant positions as real estate or debt, which are stated as separate figures. To simplify the allocation problem, life insurances and similar products are ignored.
Please enter your recommended stock portfolio share for the client at the end of the page.

Detailed description of the profile variables

- Age
  Values range 25 to 70
- Sex
  Female or male
- Married
  True or false (also stated "true" when living with permanent partner in one household)
- Kids
  Number of financially dependent children
- Education - highest educational degree
  4 categories in ascending order: high school (10years), higher education entrance qualification, professional diploma ("Berufsausbildung"), college degree

- Employment status
  3 categories: worker/employee, self-employed/business owner, retiree
- Risk preference - Motivation to take on more risk in return for higher profits
  Scale consists 5 levels from 1 to 5. 1 representing the least willingness to take on risk and 5 the most. 0 represents an unknown risk preference.
- Income
  Yearly gross income of the household. Values range € 10,000 to € 150,000
- Safe Income
  Average proportion of income that will be received with absolute certainty (excluding for example uncertain bonuses)
- Monthly spending on food and restaurants
- Housing wealth
  Values range € 0 to € 150,000,000
- Mortgage debt
  Values range € 0 to € 500,000
- Other debt
  Values range € 0 to tba
- Experience in stock investments
  3 categories: less than a year, 1 to 3 years, more than 3 years
- Investment horizon
  ALWAYS retirement savings
- Investment amount
  Total amount to be considered in the allocation. The investment amount is equal to the household's entire savings in form of financial assets. Values range € 10,000 to € 1,000,000

4.1. Questions 4.1-4.5: The portfolio experiment The financial assets are required to be allocated optimally according to the given profile. Consider the following two assets:
1. Stocks: An investment vehicle that mirrors the DAX index
2. Money market: A riskless investment into a savings account, German government bonds or other non-risky, fixed-income assets with maturities less than a year

Please state the Euro amount that you would recommend this client to invest in stocks as described above. The residual amount (investment amount - stock investment) will be allocated to the money market.

1. ______ Euro [numerical input]

4.6. Computer advice Please answer for the last 5 sections: Did you use a computer tool for your portfolio recommendation?

1. yes
2. no

4.7. Real life advice Please provide how your answer would differ had you given the advice in a personal consultation?

1. -20%
2. -15%
3. -10%
4. -5%
5. 0%
6. 5%
7. 10%
8. 15%
9. 20%

5. Final Questionnaire - Characteristics of the Participants

5.1. Sex

1. female
2. male

5.2. Married or living with permanent partner in one household

1. yes
2. no

5.3. Number of financially dependent children

1. ______ [numerical input]

5.4. Yearly gross income of the household

1. up to € 10,000
2. from € 10,000 to € 15,000
3. from € 15,000 to € 20,000
4. from € 20,000 to € 30,000
5. from € 30,000 to € 40,000
6. from € 40,000 to € 50,000
7. from € 50,000 to € 75,000
8. from € 75,000 to € 100,000
9. from € 100,000 to € 150,000
10. from € 150,000 to € 200,000
11. more than € 200,000
    [single choice]

5.5. Average proportion of income that will be received with absolute certainty (excluding for example uncertain bonuses)

1. ______

5.6. Monthly spending on food and restaurants

1. up to € 2,000
2. from € 2,000 to € 3,000

3. from € 3,000 to € 4,000
4. from € 4,000 to € 5,000
5. from € 5,000 to € 7,500
6. from € 7,500 to € 10,000
7. from € 10,000 to € 12,500
8. from € 12,500 to € 15,000
9. from € 15,000 to € 17,500
10. from € 17,500 to € 20,000
11. from € 20,000 to € 25,000
12. more than € 25,000

[single choice]//

5.7. Housing wealth

1. € 0
2. from € 0 to € 25,000
3. from € 25,000 to € 50,000
4. from € 50,000 to € 100,000
5. from € 100,000 to € 250,00
6. from € 250,000 to € 500,000
7. from € 500,000 to € 750,000
8. from € 750,000 to € 1,000,000
9. from € 1,000,000 to € 1,500,000
10. from € 1,500,000 to € 2,000,000
11. more than € 2,000,000

[single choice] //

5.8. Mortgage debt

1. € 0
2. from € 0 to € 25,000
3. from € 25,000 to € 50,000
4. from € 50,000 to € 100,000
5. from € 100,000 to € 250,00
6. from € 250,000 to € 500,000
7. from € 500,000 to € 750,000
8. from € 750,000 to € 1,000,000
9. more than € 1,000,000

[single choice] //

5.9. Saved capital

1. € 0
2. from € 0 to € 10,000
3. from € 10,000 to € 15,000
4. from € 15,000 to € 20,000

5. from € 20,000 to € 30,000
6. from € 30,000 to € 40,000
7. from € 40,000 to € 50,000
8. from € 50,000 to € 75,000
9. from € 75,000 to € 100,000
10. from € 100,00 to € 150,000
11. from € 150,000 to € 200,000
12. from € 200,000 to € 250,000
13. from € 250,000 to € 500,000
14. from € 500,000 to € 1,000,000
15. more than € 1,000,000

[single choice]//

5.10. Are you actively saving for your retirement? If yes, why?

1. No
2. Yes, I want to retire earlier.
3. Yes, I want to increase my pension.
4. Yes, because I don't want to work after my retirement to top up my pension.
5. Yes, I want to leave a reasonable heritage.
6. Other:
7. ______

[single choice] //

5.11. How much of your financial wealth should you optimally invest into risky assets?

1. ______ % of my portfolio

5.12. How much of your financial wealth is actually invest into risky assets at the moment?

1. ______ % of my portfolio

B.III. Transcript Lay Advisors

1. Age filter (*age* >= 25)

1.1. Age

1. ______ [numerical input]

2. Entrance Questionnaire - Financial Literacy

The first question group verifies that participants are able and willing to take part in the survey. Questions 1, 2, 5 and 6 are adapted from the US Health and Retirement Study (HRS) 2010 questionnaire. Questions 3 and 7 have previously been asked in the federal reserve survey on consumers and mobile financial services (2012).

Description

The next questions are about money and investments.

2.1. Safer return Imagine you are holding stocks from two different companies and you get the opportunity to add assets to your portfolio (e.g. stocks, treasury bonds, real estate). Will the risk of loosing money increase, decrease or stay the same?

1. increase
2. decrease
3. stay the same

[single choice]

2.2. High return Which asset do you think historically has paid the highest returns over a long time period, say 20 years or more – savings accounts, bonds, or stocks?

1. Saving accounts
2. Stocks
3. Bonds

[single choice]

2.3. Stock market risk If you were to invest € 1,000 in a stock mutual fund for a year, would it be possible to have less than € 1,000 when you withdraw your money.

1. Yes
0. No

2.4. Intuitiv investment behaviour If you had a larger amount of capital, could you intuitively decide how much of it you would invest in stocks or equity funds?

1. Yes
0. No

2.5. Interest First, suppose you had € 100 in a savings account and the interest rate was 2% per year. After 5 years, how much do you think you would have in the account if you left the money to grow – more than € 102, exactly € 102, or less than € 102?

1. More than € 102
2. Exactly € 102
3. Less than € 102

[single choice]

2.6. Inflation Imagine that the interest rate on your savings account was 1% per year and inflation (overall increase in prices) was 2% per year. After 1 year, would you be able to buy more than today, exactly the same as today, or less than today with the money in this account?

1. More than today
2. Exactly the same as today
3. Less than today

[single choice]

2.7. Double debt Suppose you owe € 1,000 on a loan and the interest rate you are charged with is 10% per year compounded annually. If you didn't make any payments on this loan, at this interest rate, how many years would it take for the amount you owe to double, meaning the loan reaching a value of at least € 2,000?

1. Less than 2 years
2. Between 2 and 5 years
3. 5 to 9 years
4. 10 years or more

   [single choice]

3. Expected Returns

Identical to the professional advisors' survey section 2.

4. Risk Preferences

Identical to the professional advisors' survey section 3.

5. Portfolio Recommendations

Identical to the professional advisors' survey section 4, excluding 4.6-4.7.

6. Final Questionnaire - Characteristics of the Participants

6.1. Sex

1. female
2. male

6.2. Married or living with permanent partner in one household

1. yes
2. no

6.3. Number of financially dependent children

1. ______ [numerical input]

6.4. Education

1. Highschool (10 years)
2. Highschool with higher education entrance qualification
3. Professional diploma ("Berufsausbildung")
4. Bachelor's degree
5. Master's degree
6. Doctoral degree
7. ....

   [single choice]

6.5. Employment status

1. worker/employee
2. self-employed/business owner
3. retiree
4. without engagement

[single choice]

6.6. Yearly gross income of the household

1. up to € 10,000
2. from € 10,000 to € 15,000
3. from € 15,000 to € 20,000
4. from € 20,000 to € 30,000
5. from € 30,000 to € 40,000
6. from € 40,000 to € 50,000
7. from € 50,000 to € 75,000
8. from € 75,000 to € 100,000
9. from € 100,000 to € 150,000
10. from € 150,000 to € 200,000
11. more than € 200,000

[single choice]

6.7. Average proportion of income that will be received with absolute certainty (excluding for example uncertain bonuses)

1. ______

6.8. Monthly spending on food and restaurants

1. up to € 2,000
2. from € 2,000 to € 3,000
3. from € 3,000 to € 4,000
4. from € 4,000 to € 5,000
5. from € 5,000 to € 7,500
6. from € 7,500 to € 10,000
7. from € 10,000 to € 12,500
8. from € 12,500 to € 15,000
9. from € 15,000 to € 17,500
10. from € 17,500 to € 20,000
11. from € 20,000 to € 25,000
12. more than € 25,000

[single choice]

6.9. Housing wealth

1. € 0
2. from € 0 to € 25,000
3. from € 25,000 to € 50,000
4. from € 50,000 to € 100,000
5. from € 100,000 to € 250,00
6. from € 250,000 to € 500,000
7. from € 500,000 to € 750,000
8. from € 750,000 to € 1,000,000

9. from € 1,000,000 to € 1,500,000
10. from € 1,500,000 to € 2,000,000
11. more than € 2,000,000

    [single choice]

6.10. Mortgage debt

1. € 0
2. from € 0 to € 25,000
3. from € 25,000 to € 50,000
4. from € 50,000 to € 100,000
5. from € 100,000 to € 250,00
6. from € 250,000 to € 500,000
7. from € 500,000 to € 750,000
8. from € 750,000 to € 1,000,000
9. more than € 1,000,000

    [single choice]

6.11. Saved capital

1. € 0
2. from € 0 to € 10,000
3. from € 10,000 to € 15,000
4. from € 15,000 to € 20,000
5. from € 20,000 to € 30,000
6. from € 30,000 to € 40,000
7. from € 40,000 to € 50,000
8. from € 50,000 to € 75,000
9. from € 75,000 to € 100,000
10. from € 100,00 to € 150,000
11. from € 150,000 to € 200,000
12. from € 200,000 to € 250,000
13. from € 250,000 to € 500,000
14. from € 500,000 to € 1,000,000
15. more than € 1,000,000

    [single choice]

6.12. Are you actively saving for your retirement? If yes, why?

1. No
2. Yes, I want to retire earlier.
3. Yes, I want to increase my pension.
4. Yes, because I don’t want to work after my retirement to top up my pension.
5. Yes, I want to leave a reasonable heritage.
6. Other: ______

    [single choice]

6.13. Experience in stock investments

1. less than a year
2. 1 to 3 years
3. more than 3 years

[single choice]

6.14. When was the last time you had contact to an investment consultant, asset manager or similar to talk about an investment in stocks or for your retirement?

1. Months ______
2. Years ______

6.15. How much are you willing to pay for an appointment with a financial consultant (one hour) to determine how much of your savings you should invest in stocks?

1. ______ Euro

6.16. How often do you inform yourself about news regarding investments, the capital market, etc.?

1. once a day
2. several times a day
3. weekly
4. monthly
5. less frequently than monthly
6. never

[single choice]

6.17. How much of your financial wealth should you optimally invest into risky assets?

1. ______ % of my portfolio

6.18. How much of your financial wealth is actually invest into risky assets at the moment?

1. ______ % of my portfolio

Chapter III.

In the sandbox: Can portfolio simulations help to improve efficiency in self-directed portfolios?

In the sandbox: Can portfolio simulations help to improve efficiency in self-directed portfolios?

Steffen Meyer [†] and **Matthias Rumpf**[‡]

July 7, 2020

Abstract

We analyze the log files of a web-based portfolio simulation tool for individual investors which displays any simulated portfolio next to the users' current portfolio in a risk-return diagram. We track 44,010 active investors of a German online bank between June 2014 and October 2015. 707 investors used and traded simulated positions of the tool provided by the bank. The simulation data reveals that users of the tool predominantly explore portfolios which promise higher returns, and, as a possibly unintended consequence, increase portfolio risk away from the efficient frontier. However, efficiency improvements which users achieve in terms of observable simulators' metrics, one-month expectations on return and value-at-risk, do not consistently translate into objective, long-term efficiency gains in terms of the estimated relative Sharpe ratio loss (RSRL). We conclude that designers of decision-support systems should align the information provided with objective long-term benefits, even though mitigating biases like under-diversification require a more sophisticated design.

†. University of Southern Denmark
‡. Goethe University Frankfurt

Introduction

The benefits of portfolio diversification have been well-known in finance since the introduction of Markowitz' portfolio theory (Markowitz (1952)). Nevertheless, there is abundant evidence that individual investors fail to diversify properly and thus hold inefficient portfolios. For example, investors often prefer domestic stocks over international stocks (French and Poterba (1991)), hold concentrated lottery stock portfolios (Kumar (2009)), and even ETF investors tend to underdiversify by choosing niche products (Bhattacharya et al. (2017)). Thus far, financial advice as a possible remedy has been a disappointing solution. For example because some cater to biases deteriorating even diversified portfolios (Mullainathan et al. (2012)), and others are not accepted or followed thoroughly (Bhattacharya et al. (2012)).

Digital decision support could fill this gap, given that a large share of private investors could benefit from tools that aid decision-making and trading. According to Celent (2015), a research division of Oliver Wyman, in 2015 47% of the 90 million investors in the U.S. were self-directed, and a growing number of self-directed investors, around 50%, are considered active, making at least three trades a month. Celent reports a trend towards self-directed investment, with an estimated growth rate of 4.9% compared to 1.4% for the U.S. retail investors that are not self-directed. Additionally, J.D. Power, a market research firm, reports a substantial increase in "validator" clients, from 21% to 25% over the recent past. These are mostly self-directed clients who are unwilling to delegate their investment decisions, but still appreciate on-demand access to advice, mainly in order to verify their own ideas and views.

Since the population of self-directed investors is steadily growing, computer tools that aid decision-making are a promising innovation contributing to the recent advance to "restore rationality" in consumer finance (Campbell (2016)), which still requires empirical evidence from the field. However, these tools do currently hardly exist. According to the FCA, in 2015 only 15% of UK advisory firms significantly used tools that “aid decision making and transacting” of their customers, while 46% did not use any. In computer science such tools are known as decision support systems (DSS). In consumer finance, research on DSS is scarce. For example, Looney and Hardin (2009) use an experimental website to test the design of decision support for retirement accounts to increase investor risk taking. In contrast, our research is aimed at active investors who hold risky, yet often inefficient portfolios.

We have conducted a field study together with a German online bank that has launched a portfolio simulation tool for its active, self-directed investors. It mainly consists of a simple two-dimensional risk-return diagram, i.e. a graphical display on its web-based trading platform. Upon entering the tool, clients see their current portfolio as the centered point in the risk-return diagram. The simulator calculates the expected aggregate risk and return of any arbitrary portfolio. Up to three simulated portfolios can be composed and displayed for comparison next to the current portfolio. For each position in a simulated portfolio a trading window can be opened at the click of a button. The simulation tool in our study adapts a large array of concepts suggested by current research on mitigating behavioral biases, see the next section for an overview.

It is an interactive tool that provides the opportunity to simulate and thereby experience self-created investment scenarios in a sandbox, without risking mistakes and at zero costs. Its two-dimensional display avoids information overload, while the interactive design allows active investors to discover their preferences and suitable products. We seek to explore how investors optimize portfolios and whether the digital portfolio simulation tool can indeed mitigate portfolio under-diversification.

We tracked 44,010 clients of whom 707 used the tool and traded simulated positions (called Sim-Traders); 2,521 used the tool, i.e. simulated at least three portfolios on a single day, but did not trade simulated position (Sim-Users) and 34,067 clients who did not use the simulation tool at all (Non-Users). Our data includes portfolio holdings and transactional trading data from June 2013 until October 2015. Given that the tool was introduced in June 2014, we observed 12 months prior to the introduction of the tool and 16 months post-tool introduction.

It is important to understand the distinction between two sets of portfolio efficiency metrics. First, we analyze the short-term simulator metrics visible to users of the simulation tool. Simulator metrics are the aggregated monthly expected portfolio return and value-at-risk from which we additionally derive the simulator's portfolio Sharpe ratio. The metrics are short term expectations since they are based on only six months of historical data. Second, we used the long-run relative Sharpe ratio loss (RSRL) adapted from earlier research on diversification (Calvet et al. (2007) and Gaudecker (2015)) as our main objective measure of long term portfolio efficiency. The RSRL represents the fraction of the risk adjusted benchmark return that an investor misses out due to under-diversification. We estimated the RSRL based on ex-ante expected returns using a multi-asset benchmark and the Capital Asset Pricing Model (CAPM). The RSRL dominates conventional measures of concentration such as the Herfindahl-Hirschman-Index (HHI) to evaluate portfolio diversification, by taking into account correlations in security returns and directly comparing an investor's aggregate portfolio to the mean-variance efficient benchmark. Throughout the paper, we grouped investors into quartiles according to their average pre-tool introduction RSRL in order to investigate heterogeneous treatment effects across levels of ex-ante portfolio efficiency.

Our first research question is whether the portfolio simulator leads investors to better portfolio efficiency. We test the impact of using the simulator with treatment effect estimations which exploit the panel data at hand in conjunction with conservative matching and difference-in-differences (DID) regression methods. Our estimates suggest that all investors, especially the most under-diversified, reduce portfolio risk and thereby the relative Sharpe ratio loss compared to trading without using the simulation tool. Our main model, the fixed effects regression, estimates an overall improvement of 1.6 percentage points in the relative Sharpe ratio loss, statistically significant at the 1% level. Expected portfolio standard deviation decreases by 1.07 percentage points and expected excess return declines by 0.1 percentage points. The effects are significant at the 1% and 5% level respectively. The robustness check, using a combination of matching and DID, results in estimates of similar magnitude. Nevertheless, the treatment effect estimates remain indicative, since a selection bias cannot be ruled out completely in our setup.

To assess whether investors actively optimized portfolio efficiency, when receiving the aggregate information from the simulator, we compared simulated portfolios in terms of the simulation tool metrics with the actual holdings, i.e., the starting portfolios. We determined that investors unambiguously searched for higher expected returns. Accepting higher risk, investors improved Sharpe ratios in terms of simulation tool metrics with the exception of investors in the fourth (least efficient) quartile of the RSRL distribution. On average, including only simulated portfolio with positions turned into actual trades, expected returns increased by 0.37 percentage points (from the starting portfolios' average of 1.38% p.m.). The value-at-risk increased by 0.54 percentage points (4.39), and the Sharpe ratio by 3.46 percentage points (37.96). All changes are significant at the 1% level using two-sided t-tests.

Tracing the optimization behavior induced by the simulation tool, we estimate a choice model on the trading decision for simulated security positions with multi-dimensional fixed effects. The regression of a binary trading indicator on the simulators' expected portfolio risk and return using simulated portfolio positions reveals a strong heterogeneity in preferences, especially on risk taking. Investors who hold more efficient portfolios ex-ante, show a balanced trade-off between return and risk, whereas investors holding the least efficient, risky portfolios in fact display a tendency to be attracted by both return and risk. Overall, clients' probability of trading a simulated security increases by 2.7% if expected portfolio return increases by 1% per month. If the portfolio value-at-risk increases by 1%, investors in quartiles 1 and 2 show a 1.7% and 2.2% smaller probability of trading, respectively. Investors in quartile 3 seem to be less risk averse, while investors in quartile 4 show a risk loving tendency with a positive estimate on the value-at-risk, though estimates for both Q3 and Q4 investors are insignificant.

Finally, we question whether following the simulation tool's information does indeed improve objective portfolio efficiency. We use ex-ante efficiency measures, the RSRL and the underlying expected returns and portfolio standard deviation, and analyze differences between (traded) simulations and the actual starting portfolios. A small deterioration in efficiency emerges for Q1 and Q2 investors but we find considerable efficiency gains for the less efficient investors in Q3 and Q4. Notably, starting from very high risk levels above 40% standard deviation, Q4 investors do not achieve efficiency gains by diversifying and reducing portfolio risk, but by achieving higher expected returns while accepting (or searching for) higher risk levels. While Q1 and Q2 investors improve efficiency in terms of the simulation metrics' Sharpe ratio, they reduce efficiency in terms of objective, long run ex-ante Sharpe ratio and RSRL, whereas Q4 investors achieve the opposite. We conclude that the portfolio optimization tool provides a sandbox to simulate alternatives and choose the most preferable option at zero cost. In the case of the clients holding the least efficient portfolios, this results in even stronger risk taking, while potentially ignoring salient efficiency gains. Therefore, portfolio optimization using the short-term simulator metrics does not help to mitigate biases on a long-term investment horizon. A simple scatter plot of changes in short-term versus long-term metrics of simulations over starting portfolios reveals that there is no reliable relation between the two metrics. The information provided by the simulation tool makes the resolution of under-diversification infeasible. However, given the strong preference for higher returns among all investor groups, we doubt that long-term metrics would achieve this goal. Nudging investors towards more efficient portfolios requires a more sophisticated design for a decision-support tool.

Our primary contribution is in the field of private investors' behavioral biases such as the home bias (French and Poterba (1991)), lottery stock preferences (Kumar (2009)), over-trading (Barber and Odean (2000)), or the disposition effect (Shefrin and Statman (1985)), and the growing body of literature on how to mitigate them (see, for example, the "save more tomorrow" plan by Thaler and Benartzi (2004)). User data recovered from the simulation tools' log-files offer a novel insight into the choice set and preferences of private investors with return (and risk) expectations that are exogenously provided by the tool conditional on the instrument type chosen.[1] Therefore, we also contribute to the vast literature on investor trading behavior, which includes the impact of past returns and portfolio rebalancing (see Shefrin and Statman (1985), Lakonishok and Smidt (1986), Grinblatt and Keloharju (2000), Calvet et al. (2009)) the effect of attention (see, for example, Barber and Odean (2008), da Silva Rosa and Durand (2008), and Seasholes and Wu (2007)) and investor heuristics (see, for example, Etheber et al. (2014) on moving average heuristics, on Benartzi (2001) and Bailey et al. (2011) on return extrapolation, or Jegadeesh and Titman (1993) on return chasing).

The remainder of the paper is structured as follows. In section 1, we summarize the background literature. In section 2, we describe the data as well as the simulation tool and assess who is using the tool. In section 3, we first discuss the methodology to estimate treatment effects in our setup and present the corresponding results. We estimate the overall treatment effect using monthly data to compare the portfolio efficiency before and after using the simulator (and trading) of Sim-Traders, the treated, against Non-Users, the control group. We split our investor sample into quartiles of pre-treatment averages in relative Sharpe ratio loss (RSRL) in order to investigate how investors with varying potential for efficiency improvements respond to the treatment. In section 4, we focus on the simulation data of Sim-Traders for a within-user analysis; we illustrate how simulated portfolios differ with respect to actual portfolios, and estimate preferences for risk and return among simulation users. Section 5 concludes our paper.

1. Related literature

Under-diversification is among the most common investor mistakes such as over-trading (Barber and Odean (2000)), the disposition effect (selling winning investments faster than losing ones, see Shefrin and Statman (1985)) and other special cases of under-diversification such as home bias (see French and Poterba (1991)) or lottery stock preferences (see Kumar (2009)). In the overall population of investors, under-diversification might generate only moderate return losses when tak-

1. Expected returns and VaR are based on historical return data;, therefore investors might therefore have some expectations when adding certain securities to their portfolios, but the exact precise impact on the portfolio risk and return is revealed only after the simulation.

ing the investors' investment funds and cash holdings into account (see Polkovnichenko (2005), Calvet et al. (2007)), but self-directed investors are more inclined to risky investments, as they tend to overestimate their abilities and the value of their information at hand. Overconfidence is certainly one contributory cause of the above-mentioned mistakes made by private investors. Portfolio concentration, under-diversification, and over-trading are also a consequence of other behavioral biases, primarily mental accounting and myopic investment (see, for example, Thaler et al. (1997)). Myopic behavior is related to mental accounting in the sense that investors have the impulse to frequently evaluate investments. A combination of both results in the segregation of long investment horizons into separate mental accounts for short, consecutive trading periods. This bias, also called narrow framing (Tversky and Kahneman (1985)), limits investors in their ability to consider financial decisions on an aggregate portfolio level. As a result, the benefits of diversification are either deliberately ignored by or not salient for investors. The simulation tool discussed in this article reveals promising features for ameliorating precisely such common biases. Previous literature has so far focused on educating investors about the return distribution and its riskiness in order to subsequently observe changes in risk perception and the willingness to take risks (see, for example, Kaufmann et al. (2013), Bradbury et al. (2019), or Bateman et al. (2016)).

Other measures such as financial advice, the provision of default products, as well as financial education have not yet resulted in mitigating behavioral biases to a satisfying extent (Hackethal et al. (2012), Bhattacharya et al. (2012), Campbell (2016)). Decision-support systems are thus a bold attempt to restore rationality in consumer finance. Earlier empirical research has shown many other possible approaches for improving investor decision-making, with promising results, when products or approaches have been well designed. The "save more tomorrow" plan by Thaler and Benartzi (2004) has thoughtfully addressed, used and finally mitigated behavioral biases to stimulate stock market participation. But, usually, active investors face challenges other than overcoming market entry barriers. Glaser and Weber (2007) find private investors who are unable to state even simple facts, like aggregate returns, about their own portfolios and who are often unaware of the trading cost and consequences of (under-)diversification. Counterintuitively, real trading experience does not necessarily lead to portfolio improvements (Seru et al. (2010)), while experimental evidence shows that simulated experience can increase risk-taking (Kaufmann et al. (2013)) and improve investment decisions (Bradbury et al. (2015)) by making the assets' risk distribution more salient for investors. Similarly, Goldstein et al. (2008) argue that a more interactive form of preference elicitation might help investors to construct portfolios that are more in line with their preferences.

In his extensive lecture on how to restore rational choice in consumer finance, Campbell (2016)) encourages information provision by financial service providers. Still, information disclosure requires careful design to induce optimization (Kamenica et al. (2011)). While investors must not be overloaded with information, a focus should lie on information suitable to the biases at play (Hirshleifer and Teoh (2003), Agnew and Szykman (2005)). For example graphical displays might improve investors' information perception compared to numerical displays, resulting in higher risk taking (Kaufmann et al. (2013), Glenzer et al. (2014)). Graphical information should also be provided with care, however, given that even the shape of a price path of historical returns influences investment decisions (Nolte and Schneider (2018)).

The literature on investors' trading behavior and preferences explains, to some extent, why investors hold under-diversified portfolios. Although utility functions in portfolio theory explicitly cover expected ex-ante returns and risk, research has focused primarily on the impact of past ex-post returns on investor behavior. For example, Grinblatt and Keloharju (2000) investigate the impact of past returns on trading behavior using data from the Finish Central Securities Depository. Calvet et al. (2009) expand the analysis to mutual funds and examine the portfolio rebalancing of Swedish investors after realized ex-post returns, comparing the trading patterns of winning positions compared with losing ones. They confirm the disposition effect, i.e. turnover is higher for winners than for losers (Shefrin and Statman (1985), Lakonishok and Smidt (1986)), and find that households primarily rebalance decreased risky asset shares after poor portfolio performances. Past returns of common assets are readily available in market data, but return expectations need to be either elicited from observed choices using strong model assumptions or information on them has to be requested in surveys that would have to coincide with trading. A number of studies show that investors trade stock of companies that have recently been covered in the news, which show high returns and volumes (Barber and Odean (2008), da Silva Rosa and Durand (2008), Seasholes and Wu (2007)). These attention-grabbing events make stocks appear in the choice set of investors, and it is possible that irrational beliefs about future returns increase the probability of buying assets which have shown high returns in the recent past (Jegadeesh and Titman (1993), Long et al. (1990)).

Exaggerated return beliefs might be a result of overoptimistic return extrapolation, i.e. associating current high returns with high returns the future (Benartzi (2001), Bailey et al. (2011)). This is relevant to our research, because the level of under-diversification is correlated with trend-following and overweighting of stocks with higher volatility and skewness according to Goetzmann and Kumar (2008) who further find that in the U.S. younger, less well-educated, and less-sophisticated investors show the highest level of under-diversification. There are at least two, not mutually exclusive, explanations for holding under-diversified risky portfolios. Investors might trade in mean-variance efficiency for higher skewness exposure (Mitton and Vorkink (2007)), or investors might hold (misguided) beliefs that place too optimistic probabilities on future states that deliver high returns on the purchased assets. In the optimal expectations framework (Brunnermeier et al. (2007)), investors would then optimally under-diversify given their distorted beliefs. Accordingly, we expect improvements in portfolio efficiency after using the simulation tool, if investor beliefs are indeed distorted. Even if investors voluntarily hold high skewness exposures, the simulation tool could make the trade-off between risk and return more transparent, and thus risk taking less attractive.

2. Description of the data

We work with a German brokerage firm which offers a full range of retail bank services, such as checking, term, and custody accounts as well as consumer and mortgage loans. In June 2014, the firm introduced a portfolio simulation tool as part of a risk management web service on the online banking platform. This allows its clients to back-test their own portfolio and any portfolio of their choice over a 180-day period. Simulations can be based on the clients' current portfolio positions or self-defined security watch lists. Additionally, a list of actively managed funds suggested for simulation is accessible to clients.

We prepare three data sets for our analyses. First, client data containing demographics, account characteristics, daily portfolio holdings, and trading data including the number of securities and execution prices. Second, we track simulation tool usage with time-stamped data which comprise the simulated portfolios positions with ISINs, the number of securities and euro values. Finally, our market data include daily security prices and monthly indices from Thomson Reuters Datastream, market factors small minus big (SMB) and high minus lows (HML)[2], and the 3-month Euribor time series.[3] We restrict the original sample of 113,152 clients, excluding accounts without security holdings, with less than ten months of portfolio data, and no trading activity after the introduction of the simulation tool. The resulting sample of 44,010 active clients is further split into three groups sorted by increasing use of the simulator. The first group, 34,067 Non-Users, did not use the simulator or any other component of the risk management tool. The second group, 2,521 Sim-Users, are clients who used the simulation tool with at least three simulated portfolios on a single day, but did not trade a single simulated position. This ensures that we do not pick up clients who merely entered the simulation tool, observed their own portfolios' performance, and quit the tool immediately after that. The third group, Sim-Traders, represents the 707 clients who used the simulator and traded simulated positions. All clients not sorted into these groups (but present in the overall sample) are excluded from the following analyses to obtain a clean control group of Non-Users who has neither simulated any positions nor logged on to the risk management web service.

Throughout the paper we will report our findings on the full sample and sample splits by portfolio efficiency. As explained in the next section, sample splits are based on the quartiles of the pre-tool introduction averages on the relative Sharpe ratio loss (RSRL), which is estimated based on a multi-asset benchmark using the capital asset pricing model (CAPM).

2.1. Measuring portfolio efficiency with the relative Sharpe ratio loss

We use the relative Sharpe ratio loss (RSRL) as our main measure of portfolio efficiency, which we adapt from Calvet et al. (2007) and Gaudecker (2015). The RSRL measures the loss from underdiversification taking into account correlations in security returns. The RSRL is therefore more appropriate than measures of concentration like the Herfindahl-Hirschman Index (HHI) for eval-

2. using Kenneth French's data library
3. from Bundesbank (2018)

uating portfolio efficiency and, in particular, diversification. The RSRL relies on a CAPM model for the pricing of assets on international capital markets. As Gaudecker notes, calculating expected returns directly from the securities' return time series is fraught with problems. Financial products' return histories naturally vary by their inception date, such as recently introduced exchange traded funds. This results in return time series which, potentially, cover very different market periods. Portfolio alphas inferred from this heterogeneous data would thus likely be biased conditional on the portfolio composition.

The relative Sharpe ratio loss is not based directly on historical returns. First, the correlations of portfolio securities with a benchmark are calculated, then expected returns are inferred from the benchmarks long-term return history. In contrast to conventional return expectations, this approach assumes that the correlation of asset returns with the benchmark return does not vary across time.

The most important feature of this approach for our analysis is that portfolio efficiency can be measured instantly and on any portfolio. Historical (ex-post) investment success can be measured simply by the Sharpe ratio, the portfolios' mean of historical returns divided by their standard deviation. We want to compare portfolios after the use of an optimization tool, that is, compare portfolios which, potentially, vary strongly over time and which existed only for a few days or seconds within the simulator and therefore lack a portfolio return history. Given that the relative Sharpe ratio loss uses the assets' benchmark correlations, the expected portfolio return and standard deviation, it can be calculated ex-ante. As a result, improvements in portfolio efficiency become directly visible. Furthermore, we are able to evaluate the efficiency of portfolios formed within the portfolio simulator. This means that the approach gives us a straightforward measure for comparing portfolios before simulations, simulated portfolios during optimization, and the resulting portfolios after trading.

Calculating the relative Sharpe ratio loss

The CAPM for domestic returns can be estimated using either the MSCI World Index with US$ returns minus the one month T-bill rate, $r^{\$,e}_{m,t} = r^{\$}_{m,t} - r^{\$}_{f,t}$, or a benchmark in domestic currency (euros) minus the domestic risk free rate $r^{€}_{m,t} - r^{€}_{f,t}$. Correlations are measured using all available data between the benchmark and the domestic excess return of each asset a over the domestic risk free rate $r^{€,e}_{a,t} = r^{€}_{a,t} - r^{€}_{f,t}$. We will use a multi asset benchmark with returns in euro in excess of the three-month Euribor.[4] The multi asset benchmark implements a naive diversification scheme following Jacobs et al. (2014) who claim that a 60%-25%-15% asset allocation between equity, bonds, and commodities is a robust strategy for individual investors which is not dominated by any more sophisticated portfolio optimization. The equity allocation is split equally over four regions represented by the MSCI Europe, USA, Pacific and Emerging Markets, each contributing 15% to the portfolio, resulting in an equity share of 60%. The remainder is split between bonds and

4. The Euribor time series is obtained from Deutsche Bundesbank: https://www.bundesbank.de/Navigation/DE/Statistiken/Geld_und_Kapitalmaerkte/Zinssaetze_und_Renditen/Tabellen/tabellen_zeitreihenliste.html?id=16074

commodities with 25% and 15% portfolio shares respectively. Furthermore, this benchmark, with returns in euro, is more realistic than the MSCI World from the point of view of an individual investor who is unlikely to maintain a full currency hedge on her world portfolio (Jacobs et al. (2014)). This also mitigates a large drawback of conventional CAPM models. Using only stock market risk factors, which are correlated only marginally with alternative asset classes, results in very low expected returns for these types of securities. This would exaggerate the RSRL for investors who hold a large proportion of uncorrelated but otherwise efficient securities. Nevertheless, as a robustness check we shall calculate expected returns using a Fama French 3-factor model with the unhedged world index in US$ relative to the US T-bill. Time series for the T-bill, the US size factor and the US value factor are obtained from Kenneth French's data library.[5]

We follow Calvet et al. (2007), who provide the following steps for calculating expected portfolio excess returns and variances:

1. Calculate mean $\mu_m = \overline{r_m - r_f}$ and variance $(\sigma_m)^2$ of the benchmark's excess returns
2. Estimate domestic betas β_a^D for each asset $a \in \{1, ..., N\}$ in a regression of the available historical excess returns $r_t^{€,e}$ on the benchmark's excess returns:
   $r_{a,t}^{€,e} = \beta_a \cdot r_{m,t}^{e} + \varepsilon_{a,t}$
   The $N \times N$ covariance matrix $\boldsymbol{R}$ is estimated using the regression residuals.
3. The resulting expected returns are for each asset $\mu_a = \overline{r_m - r_f}\beta_a^D$ and the covariance matrix is $\Sigma^D = (\sigma_m)\beta^D\beta^{D\prime} + \boldsymbol{R}$

The portfolios' expected return (μ_p) and standard deviation (σ_p) is determined by multiplying the expected returns and the covariance matrix by the investors' portfolio weights. With the resulting portfolio Sharpe ratio $S_p = \mu_p/\sigma_p$ and the market benchmark Sharpe ratio $S_m = \mu_b/\sigma_m$ we obtain the relative Sharpe ratio loss of the portfolio:

$$RSRL_p = 1 - \frac{S_p}{S_m} \tag{1}$$

In the CAPM framework, a perfectly diversified portfolio earns the return on the market line given a level of risk, i.e. standard deviation. Since we do not observe all risk-free assets but only risky investments, we assume that an optimal portfolio reaches the market ratio of risk and return. In this case, the relative Sharpe ratio loss would be zero. At the other extreme, if the expected return is zero, the relative Sharpe ratio loss would be 1.[6] The relative Sharpe ratio loss is hence a measure of portfolio efficiency and (under-)diversification which is convenient to interpret. It represents the fraction of the risk adjusted benchmark return that the investor misses out due to under-diversification. Put differently, it gives the level of portfolio diversification from fully diversified (RSRL=0) to no diversification (RSRL=1).

5. http://mba.tuck.dartmouth.edu/pages/faculty/ken.french/data_library.html#Benchmarks

6. In our data we even find negative expected returns which result in RSRL values greater than one. As expected, this phenomenon is especially severe when using the three-factor model instead of a multi-asset benchmark.

A related measure proposed by Calvet et al. (2007) is the return loss, which quantifies the distance of a portfolio to the market line, i.e. the distance to the maximum attainable return at a given level of risk. While the measure delivers a quantity that is easier to interpret than the relative Sharpe ratio loss, it requires knowledge of the overall portfolio risky asset share of an investor. Since we use a brokerage data set, this information is not attainable and we therefore report only the relative Sharpe ratio loss. Nevertheless, for our purpose of evaluating efficiency gains after using a portfolio optimization tool, the RSRL is perfectly suitable.

To illustrate portfolio inefficiencies, we follow Calvet et al. (2007) and Gaudecker (2015) by plotting the calculated expected returns and portfolio standard deviations in a risk-return (see Figure 1 and Figure 2). The underlying monthly data covers the treatment period 06/2014-09/2016, and is winsorized at the 1st and 99th percentile on both the ex-ante expected portfolio return, the return measure, and the ex-ante expected standard deviation, the risk measure. For Sim-Trader we consider the first month in the treatment period in which a simulated position was traded. For Non-Users we consider the first month in the treatment period in which any trade occurred. As expected, we see that both the first and the second quartile with low values of RSRL concentrate data points close the efficient frontier. Likewise, the fourth quarter shows large inefficiencies with data points moving further away. Non-Users show a larger variation between clients, illustrated by the confidence ellipses, in the first three quartiles. But the variation seems to be similar in the fourth quarter. Furthermore, the confidence ellipses reveal that, in the first three quartiles, Sim-Traders are very concentrated and close to the efficiency line, and there is almost no visible difference before and after trading. Non-Users seem to spread out more after trading, which becomes evident from expanding confident ellipses.

2.2. Investor data

Table 1 shows summary statistics on the clients and their portfolios. For each client group we report the overall average and the averages by quartiles of the pre-treatment relative Sharpe ratio loss (RSRL) distribution. The average investor in our sample is 53 years old, has been with the bank for ten years, holds 13.5 securities worth €93,400, logs on to the online platform seven times a month, generates a monthly portfolio turnover of 5% in each buys and sells (10% combined), and hold 58% of her portfolio in single stocks. It is not surprising that restricting our sample to active traders leads to an overall selection. Without restriction (not reported in the table) we observe an average portfolio value of €49,000, and the average monthly portfolio turnover is 7.4%, investors log on to their bank account on five days per month, and, on average, they allocate 37% of their portfolio value to single stocks.

When comparing the portfolio investments with other research on households, these numbers match other data sets well. For example, Barber and Odean (2001) report a turnover of 6% and Calvet et al. (2007) report an average portfolio value of US$35,000, for Swedish households. When additionally comparing our mean portfolio value to official statistics of Deutsche Bundesbank (2019) for 2017, which reports the average stock portfolio value of a German stock investor as being €43,700, and the average fund portfolio of a German fund investor as being €37,500, we find that our data appears quite representative.

We consider the (risky) securities portfolio exclusively, which makes comparison of the RSRL with earlier research difficult. Gaudecker, who uses the MSCI Europe as a benchmark, splits the Dutch household sample by its return loss quintiles, reports an average RSRL of 30% for the lowest quintile, and 88% for the highest quintile. Calvet et al. (2007) report mean RSRLs of 38% and 19% for a currency-hedged world index and an unhedged world index, respectively, using a Swedish household data set and Fama and French's 3-factor CAPM. Calvet et al. (2007) and Gaudecker (2015) have in common that they calculate the RSRL on the whole portfolio of households, and they include cash. Households with large relative cash holdings might have a large RSRL, even though they have a risky portfolio close to the mean-variance frontier. For this reason, they report numbers for the return loss, which accounts for the complete portfolio's risky share. Nevertheless, we believe that our reported average relative Sharpe ratio losses of 28.6% and 43.2% with the 90th percentiles at 55.5% and 79% (see Panel C of Table 1), for the multi asset benchmark and the 3-factor model respectively, are large but reasonable, given that we are considering active traders only.

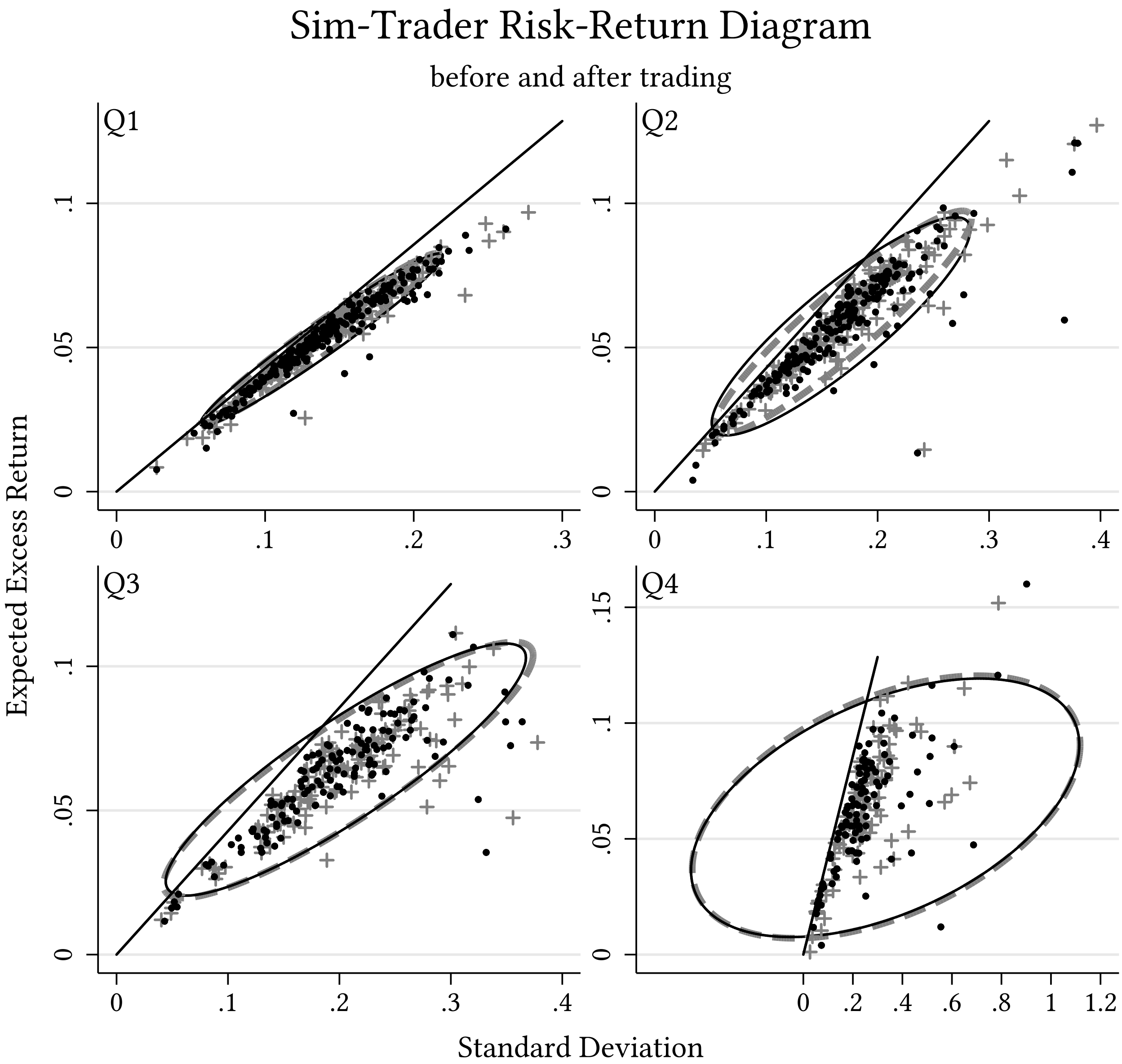


Figure 1.
Risk-return diagram by quartiles of pre-treatment RSRL averages (Sim-Traders)
Gray plus-markers (dashed ellipse) depict portfolios (95% confidence ellipse) the month before trading. Black point-markers (solid elipse) depict portfolios (95% confidence ellipse) at the end of a month in which trading took place. The underlying monthly data covers the treatment period 06/2014-09/2016, and are winsorized at the 1st and 99th percentile on both the ex-ante expected portfolio return, the return measure, and ex-ante expected standard deviation, and the risk measure. For Sim-Trader we consider the first month in the treatment period in which a simulated position was traded. For Non-Users we consider the first month in the treatment period in which any trade occurred. The solid line represents the efficient market line.

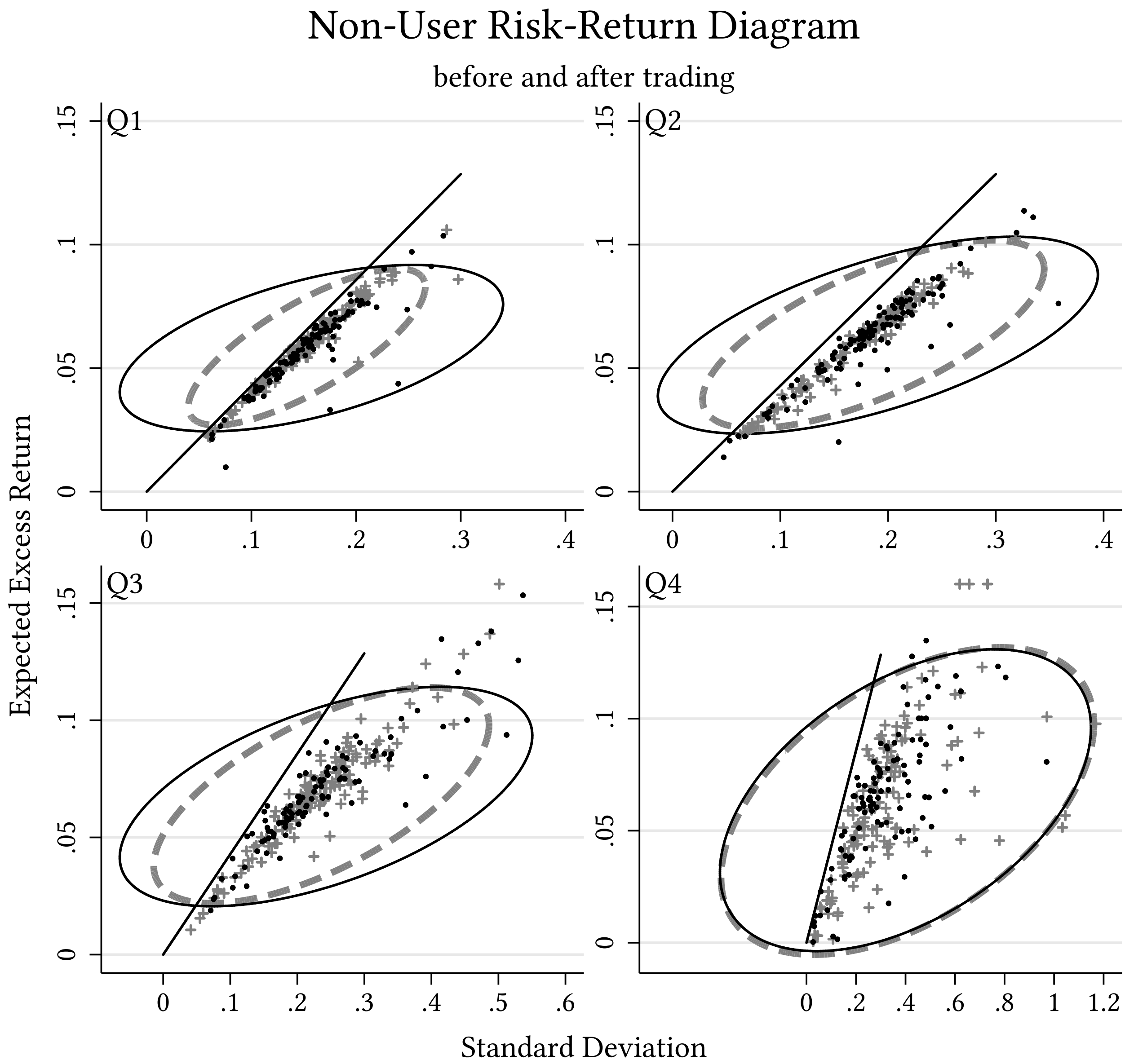


Figure 2.
Risk-return diagram by quartiles of pre-treatment RSRL averages (Non-Users)
To prevent overloading the scatter plot areas, a total of 1,000 clients are randomly selected. Confidence eclipses are calculated on the complete sample. See Figure 1 for further details.

Table 1.
Investor descriptives

This table reports summary statistics on investor demographics and characteristics (Panel A), and portfolio statistics and portfolio allocations (Panel B) by client groups and quartiles of the RSRL. The first column group (Overall) shows the mean, the 10th, and the 90th percentiles of the variable in a given row. The second column group (Non-Users) shows summary statistics for all clients with trading activity in the sample period who did not use the simulator or any other component of the risk management tool. The third column group (Sim-Users) includes clients who have used the simulation tool with at least three simulated portfolios on a single day without trading any simulated position. The last column group (Sim-Traders) represents clients who used the simulator and traded simulated positions. For each client group we show the overall mean (in column all), and the mean corresponding to the subsample of the RSRL quartile (columns Q1-Q4).

Information specific to Panel A: Relationship corresponds to the years a client has been with the brokerage. Risk tolerance is measured on a scale from 1 to 5. A Eurex customers is a client with the special privilege of trading derivatives on the Eurex exchange. Client characteristics are dummy variables equal to one if the category applies and zero otherwise. All values are as of October 2015.

Panel A: Investor demographics

	Overall			Non-Users					Sim-Users					Sim-Traders				
	Mean	P10	P90	all	Q1	Q2	Q3	Q4	all	Q1	Q2	Q3	Q4	all	Q1	Q2	Q3	Q4
Investor characteristics																		
Age	52.7	37.0	70.0	53.3	52.3	54.1	54.0	52.7	51.6	50.9	52.8	52.0	50.3	51.1	50.3	51.9	51.8	50.3
Relationship (yrs.)	10.3	8.5	16.0	10.5	10.8	10.6	10.4	10.3	9.9	10.6	9.8	9.4	9.7	9.8	10.2	10.1	9.2	9.3
Risk Tolerance	3.6	1.0	5.0	3.6	3.6	3.6	3.4	3.7	3.7	3.9	3.7	3.5	3.5	3.8	3.9	3.9	3.5	3.8
Investor characteristics (%)																		
Female	14.4	0.0	100.0	15.9	18.0	17.1	14.8	13.7	9.5	9.2	9.2	11.6	7.5	10.0	12.6	7.0	9.9	10.3
Eurex Customer	2.8	0.0	0.0	2.9	2.0	2.7	3.1	3.6	2.5	1.9	2.8	2.7	3.1	1.0	0.4	1.0	2.0	0.9
Self Employed	17.6	0.0	100.0	17.9	17.3	17.1	17.6	19.6	16.1	16.6	16.1	14.3	17.8	15.1	13.8	15.5	19.2	12.0
N	44,010			34,067	8,119	8,381	8,607	8,960	2,521	863	682	588	388	707	239	200	151	117

Table 1.
-Continued
Statistics are based on averages by client on end-of-month data over the pre-treatment period (06/2013-05/2014). Log-in days are average log-ins per month. HHI is the Herfindahl-Hirschman- Index as a measure of portfolio concentration, and the HHI 100 counts investment funds as 100 securities instead of only one to reflect the diversification effect. Turnover is measured in as a percentage relative to the overall portfolio value. Portfolio allocations are in a percentage of the total portfolio value. See Panel A's description for further details.

Panel B: Portfolio statistics

	Overall			Non-Users					Sim-Users					Sim-Traders				
	Mean	P10	P90	all	Q1	Q2	Q3	Q4	all	Q1	Q2	Q3	Q4	all	Q1	Q2	Q3	Q4
Portfolio Statistics																		
Login Days p.m.	7.0	0.3	18.9	6.0	5.7	6.4	6.2	5.7	10.8	10.3	11.8	11.3	9.8	11.8	11.4	12.8	11.5	11.4
Portfolio Value (k EUR)	93.4	4.9	195.7	93.1	138.6	105.3	77.4	55.5	104.0	140.0	117.3	70.4	51.7	90.3	119.1	90.1	74.6	52.3
Number of Securities	13.5	3.0	28.0	13.2	19.0	15.5	11.1	7.7	15.4	20.8	16.8	10.7	8.0	15.1	21.2	15.5	10.1	8.1
Portfolio Statistics (%)																		
HHI	28.5	7.1	63.7	29.4	17.1	21.4	29.1	48.5	22.7	13.9	18.7	27.6	41.7	23.2	15.1	19.4	25.6	43.1
HHI 100	19.1	0.6	51.7	19.9	3.4	10.0	22.1	42.0	13.0	2.7	8.3	19.7	34.1	13.5	2.9	8.2	19.4	36.7
Buy-Turnover	5.0	0.0	12.6	4.4	1.9	3.4	5.1	7.1	5.7	3.0	5.4	7.5	9.3	8.7	4.3	8.6	10.8	15.2
Sell-Turnover	5.0	0.0	12.0	4.6	1.9	3.7	5.4	7.1	5.2	2.7	5.1	6.8	8.2	7.7	3.4	7.8	9.4	14.4
Portfolio allocations (%)																		
Equity	78.3	35.1	100.0	78.7	78.0	80.6	82.2	74.1	76.1	73.4	76.2	82.6	72.0	76.0	73.5	76.6	81.6	73.1
Single Stocks	57.6	0.1	100.0	58.4	28.5	55.9	73.7	73.0	50.3	26.1	54.0	71.7	64.9	52.1	27.8	54.8	73.9	69.0
Equity ETF	3.9	0.0	11.1	3.6	7.2	3.8	1.6	2.0	5.2	8.7	4.7	2.5	2.5	6.0	11.4	5.1	1.7	2.3
Active Equity Funds	20.6	0.0	67.4	20.7	44.4	24.1	10.8	5.7	23.2	40.2	20.6	11.3	8.4	21.2	36.0	20.1	10.1	7.4
Fixed Income	6.5	0.0	23.1	6.3	6.5	5.7	5.2	7.6	7.8	7.5	7.3	6.3	11.8	7.0	7.9	6.5	4.4	9.1
Bonds	3.5	0.0	6.6	3.5	2.2	3.1	3.4	5.2	3.4	1.9	3.8	3.2	6.6	3.1	2.2	2.6	2.8	6.1
FixInc ETF	0.5	0.0	0.0	0.4	0.6	0.3	0.2	0.4	0.9	1.2	0.7	0.3	1.3	0.9	1.3	1.1	0.4	0.2
Active FixInc Funds	2.5	0.0	7.0	2.4	3.7	2.3	1.6	2.0	3.5	4.3	2.8	2.7	3.9	3.0	4.3	2.9	1.2	2.8
Balanced Funds	4.4	0.0	15.5	4.0	7.1	4.5	3.0	1.7	7.3	11.1	7.7	3.9	3.1	6.9	10.1	6.7	4.8	3.1
Other	5.0	0.0	15.4	5.0	4.2	4.3	4.3	7.1	4.2	4.4	4.0	3.0	6.3	4.9	4.9	4.9	3.4	6.7
N	44,010			34,067	8,119	8,381	8,607	8,960	2,521	863	682	588	388	707	239	200	151	117

Table 1.
-Continued

Statistics are based on averages by client on end-of-month data over the pre-treatment period (06/2013-05/2014). Log-in days are average log-ins per month. This panel reports summary statistics on portfolio efficiency based on a CAPM model using a multi-asset benchmark for the first set of rows (see *Multi-Asset Efficiency*) and a 3-Factor Model (see *3-Factor Model Efficiency*). See section 2.1 for details of how the efficiency measures are calculated.

Panel C: Portfolio efficiency

	Overall			Non-Users					Sim-Users					Sim-Traders				
	Mean	P10	P90	all	Q1	Q2	Q3	Q4	all	Q1	Q2	Q3	Q4	all	Q1	Q2	Q3	Q4
Multi-Asset Efficiency																		
Rel. Sharp R. Loss	28.6	8.9	55.5	29.4	9.6	18.9	30.2	56.4	23.4	9.1	18.6	29.8	53.7	23.8	9.5	18.9	29.8	53.6
Exp. Excess Return	6.3	3.4	8.9	6.3	5.8	6.4	6.8	6.4	5.8	5.4	5.8	6.5	5.7	5.9	5.4	6.0	6.5	5.9
Exp. Standard Dev.	24.1	10.7	38.0	24.7	15.1	18.4	23.2	40.9	19.8	14.0	16.7	22.2	34.3	20.3	14.1	17.3	22.1	35.8
Sharpe Ratio	30.6	19.1	39.0	30.2	38.7	34.7	29.9	18.7	32.8	38.9	34.9	30.1	19.8	32.7	38.8	34.8	30.1	19.9
3-Factor Model Efficiency																		
Rel. Sharp R. Loss	43.2	16.0	79.2	43.8	16.4	31.5	46.8	78.7	39.8	16.9	31.5	46.1	74.8	40.2	17.0	31.8	46.6	74.1
Exp. Excess Return	4.7	1.5	8.3	4.8	6.8	5.3	4.5	2.7	4.3	6.1	4.6	3.8	2.3	4.4	6.1	4.5	4.0	2.5
Exp. Standard Dev.	23.9	10.4	37.7	24.6	20.1	19.1	21.5	36.8	19.6	18.2	16.7	17.9	28.1	20.0	18.0	16.5	19.6	28.2
Sharpe Ratio	22.9	8.4	33.9	22.7	33.8	27.7	21.5	8.6	24.3	33.6	27.7	21.8	10.2	24.2	33.5	27.5	21.6	10.5
N	44,008			34,067	8,477	8,323	8,365	8,902	2,521	651	713	681	476	706	178	209	180	139

2.3. Portfolio Simulation Data

The simulation tool was introduced in June 2014 to allow clients to back-test their own and any arbitrary portfolio over a 180-day period based on their current portfolio positions or self-defined alternative portfolios. The tool targets insufficient knowledge of aggregate portfolio information by providing a simple environment which helps to evaluate investments in the context of the clients' entire portfolios. It serves investors challenged by the trade-off between the risk and return of different products by visualizing efficiency gains between simulated portfolios. The optimizer generates a graphical display by plotting a representative dot in a risk-return diagram for each simulated portfolio, marking its expected return on the y-axis, and its value-at-risk (VaR) on the x-axis. Figure 3 provides an illustration of the tool's output, which appears prominently at the top of the simulation tool's page on the banks' online platform. Up to four portfolios, three simulation results plus the actual user portfolio can be compared on the two-dimensional plane. Both risk and return values are one-month expectations, based on historical data over the past six months. The VaR is provided as a percentage loss and calculated at the 5% level.

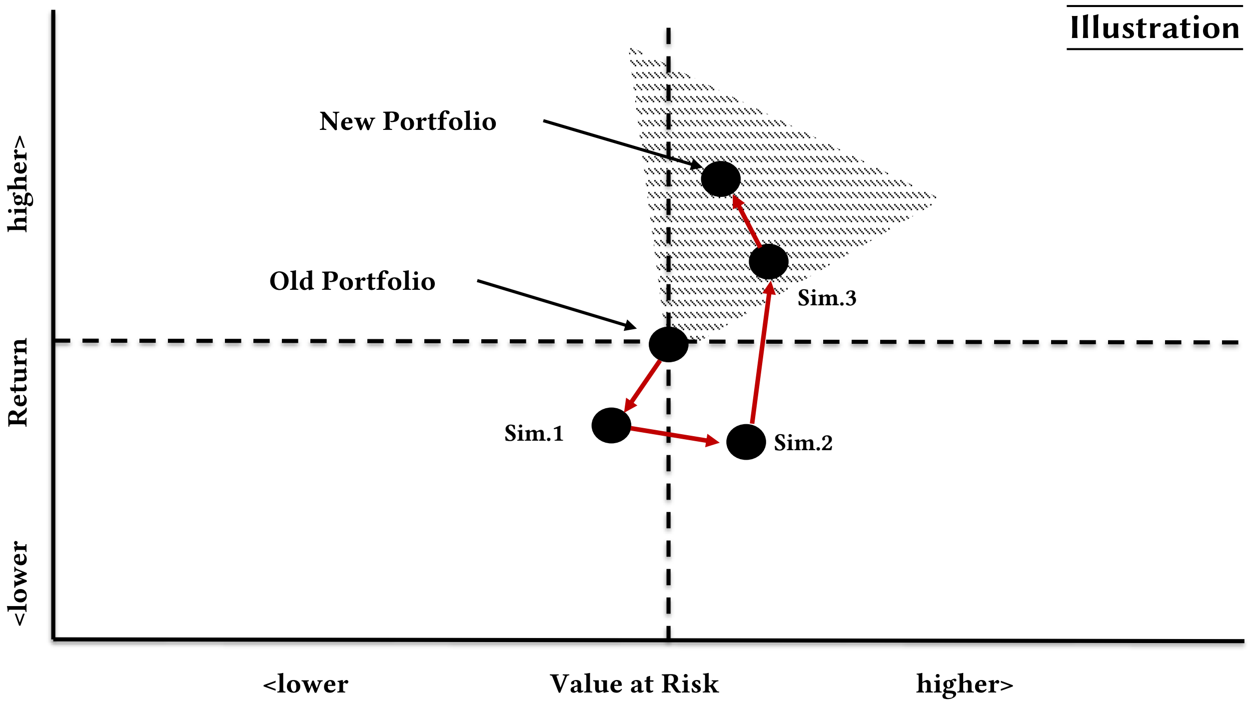


Figure 3.
Graphical output of the portfolio simulation tool
The figure illustrates how simulated portfolios are displayed in the simulation tool's risk-return diagram and shows a potential path over various simulation runs. The shaded area represents a common area for simulated portfolios, i.e. simulated portfolios show higher risk but also higher returns in comparison to the actual portfolio.

We use the simulation data to evaluate two main questions. The first question is how do investors optimize their portfolios? This is achieved by analyzing simulations which are based on the clients' actual portfolios. We look at the differences between simulations and portfolios with respect to the performance and efficiency measures observable within the simulator. The second question is what drives the trading decision for securities tested during the portfolio optimization process. To this end, we generate two data sets for portfolio simulations. First, we match clients' actual portfolios to the simulated portfolios. Second, based on the security position differences between simulations and actual portfolios, we match simulated positions to the clients' trades that occurred up to ten days after a simulation day.

Simulated portfolios and actual portfolios

In order to follow the optimization strategy of investors, we disregard simulations that are not realistically related to clients' actual portfolios: We drop all simulations that are identical to the actual portfolio and simulations with the following features which identify simulated portfolios not related to clients' actual portfolios: Simulations with more than five times the number of distinct securities for investors holding more than five securities. Simulations containing less than one-fifth of the actual portfolio's security count for investors holding more than ten securities, simulations with more than five times the actual portfolio's value or less than one-fifth of the portfolio value for investors holding €5,000 or more in total portfolio value, and simulations which include less than one-fifth of the actual portfolio's securities for investors holding more than five securities. Our data set provides the values on performance and VaR from the simulation tool for active simulations as well as actual portfolios. We drop about 2% of simulations that have expected returns of exactly zero for either portfolios or simulations, given that these indicate insufficient data for performance calculations in the bank brokerage's security time series data base.

Table 2 reports the distributions of distinct simulation counts and simulation days across Sim-Users and Sim-Traders in Panel A and B, respectively. Panel C reports the distribution of simulated and traded simulation positions across Sim-Traders. We notice that Sim-Traders used the simulator much more frequently than Sim-Users. While Sim-Users ran an average of 11 simulations on 1.5 distinct days, Sim-Traders ran 46 simulations on 4.6 days during the treatment period (see Panel A and B). Concentrating on Sim-Traders' simulations, we see that, on average, each simulation contained 2.3 buy and 2.1 sell positions. Buy positions entered simulations slightly more frequently: we count around 76,500 buy positions and only around 75,000 sell positions in around 32,000 simulations. Finally, while 17% of simulated buy positions were traded, only 13.0% of simulated sell positions were actually sold.

Simulated positions and actual trades

Comparing the buy positions of Non-Users and Sim-Traders, we see that Sim-Traders trade much more frequently, but in smaller positions. For example, Non-Users traded around 18 equity positions with an average value of €5,362 during the treatment period (see columns 2 and 3 in Table 3), whereas Sim-Traders bought around 42 equity positions with an average value of €4,237. In comparison, simulated positions that were actually traded are, on average, two to three times larger than the average traded positions of Sim-Traders (compare columns 6 and 12). Hence, the question arises as to whether clients seek help by using the simulator when positions are larger

Table 2.
Count of Simulations and Simulated Positions
Panel A shows the distributional statistics and distinct counts of simulation scenarios and simulation days for Sim-Users. Panel B is identical to Panel A, except for showing data on Sim-Traders. Panel C shows the count of securities positions (and traded positions) on data aggregated by simulations of Sim-Traders only.

Panel A: Simulation count by Sim-User

	N	investors	mean	p5	p10	p25	p50	p75	p90	p95	max
Simulation Scenarios	20 563	1891	10.87	1	1	2	5	11	23	38	840
Simulation Days	2835	1891	1.50	1	1	1	1	2	2	3	61

Panel B: Simulation count by Sim-Trader

	N	investors	mean	p5	p10	p25	p50	p75	p90	p95	max
Simulation Scenarios	30 115	655	45.98	2	2	6	15	43	107	171	1466
Simulation Days	3015	655	4.60	1	1	1	2	4	9	15	207

Panel C: Position count by Sim-Trader simulation

	N	simulations	mean	p5	p10	p25	p50	p75	p90	p95	max
Buy Positions	76 546	33 056	2.32	0	0	1	1	3	5	7	94
Traded	12 854	33 056	0.39	0	0	0	0	1	1	2	13
Sell Positions	74 912	33 056	2.27	0	0	0	0	2	6	11	145
Traded	9722	33 056	0.29	0	0	0	0	0	1	2	12

and thus have greater impact on the overall portfolio.

In both trades and simulations we see the current shift towards low-cost products. Portfolios of Sim-Traders in the pre-treatment period consisted to 6.0% of equity ETF and 21.2% of actively managed equity funds (see Table 1, Panel B). Trades during the treatment period consist of 10% equity ETF and 11.6% active equity funds. Interestingly, the share of active equity funds is much higher in simulations and in traded simulation positions, i.e. 20% and 15%, respectively. This is most likely due to the fact that, within the portfolio simulator, actively managed funds are suggested as a default investment opportunity next to the possibility of entering ISINs autonomously. This would also be reflected in the relatively large share of active fixed income funds of 8.6% in simulations (see column 7 in Table 3). Non-Users trade only 7.3% in active equity and 1.3% in active fixed income funds. Finally, Non-Users' and Sim-Traders' asset class distributions of trades differ mainly in the proportion of other assets, which is very large for Non-Users with 18%. Other assets include commodities, real-estate, and money market investments, but also certificates and derivatives that could invest into any asset class. In summary, the descriptive statistics on trades and simulations do not reveal an obvious optimization effect on the trading behavior of Sim-Traders, such as a drastic reduction of single stocks. Nevertheless, suggested investments like the pre-set of actively managed funds seem to enter the choice set of clients.

Table 3.
Asset class distribution of trades and simulations (buy positions only)
The table shows the asset class distributions of trades and simulations. All values are aggregated over the treatment period (06/2014-10/2015). "Share" is the corresponding asset class's percentage of total value traded (or simulated) across all investors. "Count" is the average number of traded (simulated) securities per asset class and investor. "Value" is the average position value. In the case of simulations, only simulations that are related to the investors' actual portfolios are included.

	Non-Users			Sim-Traders								
	Trades			Trades			Simulations			Traded Sim. Positions		
	Share	Count	Value	Share	Count	Value	Share	Count	Value	Share	Count	Value
Equity	74.7	17.6	5362.3	78.4	42.4	4237.1	77.2	86.1	9062.1	80.5	3.2	11 073.6
Single Stocks	60.3	15.6	4893.7	56.6	35.6	3746.4	47.8	54.4	9157.0	53.3	2.2	10 333.6
Equity ETF	7.0	0.9	7400.3	10.1	3.5	5629.5	9.7	14.2	7732.4	12.0	0.5	13 602.6
Active Equity Funds	7.3	1.1	6228.8	11.6	3.4	4854.8	19.8	17.5	6816.4	15.3	0.6	11 567.7
Fixed Income	4.3	0.7	9585.0	5.4	1.3	5847.1	12.8	10.4	7883.2	6.1	0.2	18 059.3
Bonds	2.2	0.4	11 380.9	1.2	0.3	5111.9	0.3	0.1	8850.4	0.6	0.0	16 091.7
FixInc ETF	0.7	0.1	8734.2	2.1	0.4	6273.8	3.9	3.6	9978.7	2.3	0.1	17 208.5
Active FixInc Funds	1.3	0.3	7762.6	2.1	0.6	6390.6	8.6	6.6	7075.8	3.1	0.1	16 853.2
Balanced Funds	3.1	0.3	7829.3	4.3	1.1	7251.9	3.9	4.1	11 173.0	5.6	0.2	21 415.3
Other	17.9	9.2	25 445.1	11.9	13.0	46 899.3	5.4	4.6	10 462.9	7.8	0.2	15 004.6
N investors	24989	24989	24989	631	631	631	602	602	602	504	504	504

2.4. Who uses the simulator?

Non-Users are almost uniformly distributed across RSRL quartiles, which is only natural given that they represent almost 80% of the whole sample (see Table 1). Sim-Traders, on the other hand, are recruited to 34% from the first quartile and only to 17% from the last quartile, meaning that, overall, Sim-Traders invest more efficiently than the sample of active traders even before the treatment. We explore further selection factors in logit models that regress, for example, a dummy variable equal to 1 for Sim-Users and equal to 0 for Non-Users on a set of pre-treatment client and portfolio characteristics; see Table 4.

All effects for Non-Users selecting into Sim-Users (column 1) compared to Sim-Traders (column 2) go in the same direction but with larger magnitudes, except for the effects on Eurex customers, turnover, and the single stock portfolio allocation. Most interestingly, higher turnover decreases the chance of switching from Non-Users to Sim-Users, but increases the chance of switching to Sim-Traders. This suggests that demographics are more important if trading is not the motive which induced using the simulator.

Using the overall samples interdecile range (IDR),[7] we infer that average log-in days per month is the most important factor for selecting into using the simulator with a 10.75% higher probability (on a change in the login days equal to the IDR=18.5). Age decreases the probability by 2.5% (IDR=33) and the relationship length by 3% (7.5). The selection effect of portfolio efficiency is also large with -5.2% (5.5%) for expected portfolio returns, 6.8% (27%) for the portfolio standard deviation, and -6.6% (46.6%) for the RSRL. All other factors have an effect size of less than 1%. It is important to note that efficiency measures are strongly interdependent and effects are estimated ceteris paribus. Judging by the relative Sharpe ratio loss, investors with less diversified portfolios are ceteris paribus more likely to use the simulator. Holding efficiency fixed, higher expected returns decrease the probability of switching to Sim-Users and higher portfolio risk, i.e. standard deviation, increases such a probability. It seems that more sophisticated investors are less likely to use the decision-support system, confirming our initial observation on the summary statistics.

As discussed, selection into Sim-Traders is very similar, only few characteristics have an economically significant impact of switching between Sim-Users to Sim-Traders (see column 3). For example, Eurex Customers are 19% less likely to switch. An IDR increase results in a change in probability of 4.0% (given IDR=18.6) for log-in days, 3.1% (12%) for turnover, -2.2% (33) for age, 2.1% (100%) for single stocks, and -1.8% (46.6%) for RSRL. We thus infer that older, more active investors holding more efficient portfolios are more likely not only to test but also to trade simulated positions. The same holds true for switching from Non-Users to Sim-Users with larger effect size and a shorter relationship to the bank as an additional significant factor.

7. the distance between the 9th and the 1st decile

Table 4.
Logit regression for client group selection
This table reports the marginal effects from a logit regression. The dependent variable (y) is binary. For example, the model in column 1 shows regression coefficients with y=1 for simulation users who did not trade simulated positions (Sim-Users) and y=0 for actively trading clients traders who did not log on to the risk management tool at all (Non-Users). Independent variables are pre-treatment averages, except for the time-invariant variables age, the relationship length, the dummy for Eurex trading and self-employment. Portfolio value is on the log scale. Turnover, portfolio shares, and efficiency measures are averages over the pre-treatment period and their values enter the regression unscaled, for example 1% turnover is 0.01, a single stock share of 55% enters as 0.55, and an excess return of 3.1% enters as 0.031. Together with marginal effects as a percentage this increases the information provided with a limited amount of decimals. P-values in parentheses are based on heteroscedasticity robust z-statistics. ***, **, and * denote statistical significance at the 1%, 5%, and 10% levels, respectively.

	(1)	(2)	(3)
y=1	Sim-Users	Sim-Traders	Sim-Traders
y=0	Non-Users	Non-Users	Sim-Users
Age	−0.075***	−0.031***	−0.066***
	(0.000)	(0.000)	(0.000)
Relationship (yrs.)	−0.409***	−0.156***	−0.030***
	(0.000)	(0.000)	(0.000)
Eurex Customer	−1.530***	−2.378***	−19.051***
	(0.000)	(0.000)	(0.000)
Self Employed	−0.651***	−0.362***	−0.928***
	(0.000)	(0.000)	(0.000)
Login Days p.m.	0.578***	0.207***	0.213***
	(0.000)	(0.000)	(0.000)
Number of Securities	−0.012***	−0.007***	−0.008***
	(0.000)	(0.000)	(0.000)
Log Portfolio Value	−0.315***	−0.174***	−0.526***
	(0.000)	(0.000)	(0.000)
Turnover	−1.992***	1.102***	25.506***
	(0.000)	(0.000)	(0.000)
Single Stocks	0.543***	0.604***	2.125***
	(0.000)	(0.000)	(0.000)
Equity ETF	2.746***	1.943***	9.570***
	(0.000)	(0.000)	(0.000)
Bonds	−3.552***	−1.265***	−1.058***
	(0.000)	(0.000)	(0.000)
FixInc ETF	7.573***	2.192***	−5.217***
	(0.000)	(0.000)	(0.000)
Exp. Excess Return	−94.445***	−32.166***	−17.361***
	(0.000)	(0.000)	(0.000)
Exp. Standard Dev.	2.775***	0.965***	2.250***
	(0.000)	(0.000)	(0.000)
Rel. Sharp R. Loss	−14.192***	−5.094***	−3.798***
	(0.000)	(0.000)	(0.000)
Pseudo R2	0.092	0.098	0.013
Observations	36,586	34,772	3,228

3. Does using the simulator improve portfolio efficiency?

3.1. Methodology

Regression specifications

In this section we carefully derive regression tools appropriate for the observational data at hand. We have to deal with the fact that the treatments of using simulation tool were not randomly assigned within the investor population. While investors were invited by mail (physical and electronic) to use the risk management web service, the invitations do not serve as a valid instrument for using a single specific component like the simulator. A single instrument is insufficient when multiple treatments are potentially in play. Observable investor characteristics such as age and portfolio value, as well as unobservables such as individual habits, intrinsic motivation, or social interactions, are likely to drive the assignment to the treatment. Hence, we must discuss methods for causal inference in the presence of selection bias which impose assumptions likely to hold in our setup. We will derive and discuss three models.

As our first and primary benchmark model we will use a panel extension to a difference-in-difference (DID) framework adopted from Wooldridge (2012). The second model is a combination of matching and symmetric DID, as suggested by Chabé-Ferret (2017), and the third model includes information on lagged outcome variables in a sequential matching approach (see also Wooldridge (2012)).

Introduction to causal inference in the presence of selection bias

We start with a simple regression model that measures the mean difference τ in outcome Y_i between the treated, $T_i = 1$, and control group, $T_i = 0$, to formally illustrate the role of selection (see Angrist and Pischke (2008), ch. 3.2).

$$Y_i = \alpha + \tau T_i + \eta_i \tag{2}$$

The OLS estimate of τ is consistent if the assignment to the treatment T_i is uncorrelated with the outcome, i.e. the residual η_i. But the decision to use the simulator is likely to depend on observable investor characteristics, $\boldsymbol{X}_i$, such as portfolio value or investment experience, which are part of the residual in equation (2):

$$\eta_i = \boldsymbol{X}_i'\delta + \nu_i \tag{3}$$

Hence τ, in equation (2), is prone to an omitted variable bias since exogeneity of T_i with respect to η_i is violated. The bias can be removed by adding all factors that influence the treatment to the regression:

$$Y_i = \alpha + \tau_{cond} T_i + \boldsymbol{X}_i'\delta + \nu_i \tag{4}$$

Using observable investor characteristics $\boldsymbol{X}_i$, τ_{cond} is the causal treatment effect if and only if Y_i and T_i are independent conditional on $\boldsymbol{X}_i$, i.e. the conditional independence assumption (CIA) holds. Intuitively, this condition is also called 'selection on observables',[8] given that the selection into the treatment is fully captured by the observables $\boldsymbol{X}_i$.

Difference-in-differences

If selection on observables does not hold, a difference-in-differences (DID) approach might help. DID estimates control for selection attributable to time-constant factors such as investor preferences and characteristics. By differencing outcomes over at least two time periods, time-constant investor fixed effect are completely removed. Following Angrist and Pischke (2008) we extend the model in equation (4) by adding time-fixed effects λ_t and a vector of unobserved investor fixed effects A_i. We use W_{it} as the treatment dummy variable, which is equal to one if individual i has been treated in period t and zero otherwise.

$$Y_{it}(T_i, W_{it}) = \alpha + \lambda_t + \tau_{DID} W_{it} + \boldsymbol{X}_i' \delta + A_i' \gamma + \varepsilon_{it} \tag{5}$$

The mechanics behind DID become clear when considering only a pre-treatment (t=1) and a post-treatment period (t=2), ignoring the covariates X for the moment. Note that, in the first period, the treatment dummy is equal to zero for both groups, $W_{i1} = 0$. First, we difference over the time dimension, $t = 1$ minus $t = 0$. For the treated, with $T_i = 1$ and $W_{i2} = 1$, the outcome equation is

$$Y_{i1}(T_i = 1, W_{i1} = 1) = \alpha + \lambda_1 + \tau_{DID} \cdot 0 + \boldsymbol{A}_i' \gamma + \varepsilon_{i1} \tag{6}$$

$$Y_{i2}(T_i = 1, W_{i2} = 1) = \alpha + \lambda_2 + \tau_{DID} \cdot 1 + \boldsymbol{A}_i' \gamma + \varepsilon_{i2} \tag{7}$$

$$\Delta Y_i(T_i = 1) = Y_{i2}(T_i = 1, W_{i2} = 1) - Y_{i1}(T_i = 1, W_{i1} = 0) \tag{8}$$

$$= \lambda_{2-1} + \tau_{DID} + \varepsilon_{i,2} - \varepsilon_{i,1} \tag{9}$$

Similarly for the control group with $W_{j2} = 0$ (using subscript j for clarity), we have

$$Y_{j1}(T_j = 0, T_{j1} = 0) = \alpha + \lambda_1 + \tau_{DID} \cdot 0 + \boldsymbol{A}_j' \gamma + \varepsilon_{j1} \tag{10}$$

$$Y_{j2}(T_j = 0, T_{j2} = 0) = \alpha + \lambda_2 + \tau_{DID} \cdot 0 + \boldsymbol{A}_j' \gamma + \varepsilon_{j2} \tag{11}$$

$$\Delta Y_j(T_j = 1) = Y_{j2}(T_j = 0, T_{j2} = 0) - Y_{j1}(T_j = 0, T_{j1} = 0) \tag{12}$$

$$= \lambda_{2-1} + \varepsilon_{j,2} - \varepsilon_{j,1} \tag{13}$$

With $\lambda_{2-1} = \lambda_2 - \lambda_1$, and zero mean residuals, the second differencing, treatment - control, yields the DID treatment effect after taking expectations:

$$\Delta Y_i(T_i = 1) - \Delta Y_j(T_j = 0) = \lambda_{2-1} + \tau_{DID} + \varepsilon_{i,2} - \varepsilon_{i,1} - (\lambda_{2-1} + \varepsilon_{j,2} - \varepsilon_{j,1}) \tag{14}$$

$$E(\Delta Y_i(T_i = 1) - \Delta Y_j(T_j = 0)) = \tau_{DID} \tag{15}$$

8. Other labels for this assumption include unconfoundedness, exogeneity, ignorability (see, for example, Imbens and Wooldridge (2009)).

Here, we took advantage of two crucial assumptions. First, bias stability (Lechner (2011)), requires selection on unobservables, i.e. $\gamma_t \boldsymbol{A}_i$, to be time-constant: $\gamma_t = \gamma$, and all elements of A_i remaining fixed over time.[9] Second, at the heart of the DID approach lies the common (or parallel) trends assumption (Angrist and Pischke (2008), Lechner (2011)), which means that changes in the outcome Y_i would be equal for both treated and control, had the treated group not been treated:

$$(Y_{i2}(T_i = 1, W_{i2} = 0) - Y_{i1}(T_i = 1, W_{i1} = 0)) \tag{16}$$

$$=(Y_{j2}(T_j = 0, W_{j2} = 0) - Y_{j1}(T_j = 0, W_{j1} = 0)) \tag{17}$$

In terms of the regression equation (4), this translates into the time effects λ_{2-1} being equal for the treated and control group.

Finally, when the common trend assumption does not hold, we can introduce available control variables $\boldsymbol{X}_i$, that must be exogenous with respect to the treatment and, if necessary, individual time trends $\kappa_i \cdot t$, to meet the parallel trends assumption conditional on investor specific controls (Lechner (2011), Angrist and Pischke (2015)).

Model 1: Fixed Effects (FE) Panel Regression

We extend the DID framework to fully exploit multiple time periods in the case of arbitrary treatment periods, meaning treatment periods that can start and end at different (and possibly multiple) times across investors i. This results in a series of treatment dummies, $\boldsymbol{W}_i = (W_{i1}, .., W_{iT})$. We implement the following fixed effects panel regression suggested by Wooldridge (2012) (p. 102), which will serve as our main model.[10]

$$Y_{it} = \tau_{FE} W_{it} + \sum_{i=1}^{I} \beta_i \cdot 1_{I_i=i} + \sum_{t=1}^{T} \lambda_t \cdot 1_{T_i=t} + \boldsymbol{X}'_{it} \boldsymbol{\delta}_1 + W_{it} \cdot (\boldsymbol{X}_{it} - \xi_t)\, \boldsymbol{\delta}_2 + e_{it} \tag{18}$$

Here α, $\sum_{i=1}^{I} \beta_i \cdot 1_{I_i=i}$ represent a set of individual fixed effects, $\sum_{t=1}^{T} \lambda_t \cdot 1_{T_i=t}$ a set of time effects, and ξ_t the period-specific population average, of time varying covariates $\boldsymbol{X}_{it}$.

This setup assumes unconfoundedness of the sequence of treatments, conditional on constant unobserved heterogeneity, captured by fixed effects, β_i, and a history of covariates, $\boldsymbol{X}_{it}$. Importantly, $\boldsymbol{X}_{it}$ is not allowed to include factors that react to shocks to previous outcomes, e.g. no lagged outcomes (see Wooldridge (2012)). In line with this, $\boldsymbol{X}_{it}$, must be exogenous with respect to the treatment $\boldsymbol{W}_i$, i.e. the treatment must have no causal effect on the observed covariates, which is known as the exogeneity assumption (see e.g. Lechner (2011)).

9. Due to the second differencing γ could even be time-varying as long as the unobservables were be completely balanced between the treated and control group, as discussed in O'Neill et al. (2016). This is an assumption that cannot be tested. Combining DID with matching is partly motivated by hoping to balance endogenous unobserved variables with time-varying coefficients, and thus reducing bias in the DID estimator (see also Lechner (2011)). Given that our pre-treament and post-treament periods span only a few months, we argue that effects on unobserved investor characteristics are unlikely to vary over the observation period.

10. The regression equation is adapted from Angrist and Pischke (2008), Angrist and Pischke (2015), and Imbens and Wooldridge (2009).

We estimate this model 1, equation (18), using the STATA command *reghdfe* (high dimensional fixed effects regression), which absorbs fixed effects over the time and cross-sectional dimensions. P-values are based on clustered standard errors at the investor level to cope with serial correlation.[11]

Finally, we must emphasize that identification of model 1 relies on the assumption that selection into the treatment can be entirely attributed to observable covariates and, most importantly, to unobservable fixed effects. The latter are controlled for by absorbing investor fixed effects.

In our data set practically all time varying control variables are potentially influenced by the treatment. For example, past experiences with the simulator might influence investor behavior measured by portfolio turnover, risk taking or platform login frequency. Investor characteristics such as age are considered as constant, since we only analyze a short time period. The only likely candidate for an exogenous control variable is a monthly indicator for informational mailings about the risk management service that were frequently sent to the clients. Time-constant investor characteristics that might have an influence on treatment participation are captured by the investor fixed effects and are therefore omitted. Investor fixed effects also control for pre-treatment averages of control variables that are frequently used to capture selection on observables while preserving exogeneity.[12] The fixed effects model has one considerable disadvantage. Using a long panel data set, it is very likely that there is a feedback effect between the treatment and the outcome variables. The treatment might induce changes to outcomes in period t=1, which, in turn, influences the selection into the treatment in future periods. This dynamic can be accounted for in a sequential matching approach, also suggested in Wooldridge (2012), which we will discuss for the third model.

Matching

The DID estimator might be biased if selection is due to unobserved time-varying factors, such as information shocks or a large loss suffered due to strong portfolio concentration. The common trends assumption might not hold, for example, if investors with a recent financial loss are more likely to seek help using the risk management tool and, at the same time, are more likely to profit from using it. This would create a dis-balance in the potential treatment effect between the treated group and the control group. In this case, we have to rely on observable variables to selection. But even if selection is fully accounted for by observables, the estimation of τ_{cond}, equation (4), might not be feasible if the treatment and control groups are entirely different in terms of $\boldsymbol{X}_i$. This becomes more clear when thinking about a manual (non-regression) estimator of τ_{cond}. If observations (investors) are grouped by values of $\boldsymbol{X}_i$, τ_{cond} can be estimated by averaging group differences in the outcome Y_i between the treated and control groups, weighted by the marginal distribution of $\boldsymbol{X}_i$ (Angrist and Pischke (2015), p.42). If treated and controls are very distinct, they cannot be matched in terms of $\boldsymbol{X}_i$ and differences in the outcomes between the two

11. See Bertrand et al. (2004) for an extensive discussion on standard error estimation for DID models.

12. For example Wooldridge advises against matching on the control variables' first periods see Wooldridge (2012), Imbens and Wooldridge (2009).

groups cannot be calculated.

Matching solves this problem by improving the overlap in covariate distributions (see, for example, Imbens and Wooldridge (2009) for a recent discussion). In a simple form, the matching approach is implemented by searching, for each individual in the treated population, a control group member who is an exact match in terms of $\boldsymbol{X}_i$. All unmatched individuals are discarded. The resulting estimator averages differences in the outcomes $Y_i(T_i)$ between groups which are equal in terms of observable characteristics (see Abadie and Imbens (2006)):

$$\hat{\tau}_{match} = \frac{1}{N} \sum_{N}^{i=1} \left(\hat{Y}_i(1) - \hat{Y}_i(0) \right) \tag{19}$$

It is important to note that both regression and the matching estimator rely on T_i being conditionally independent of the Y_i (v_i), i.e. the CIA to hold. In fact, matching is often complemented by regression in order to obtain more efficient standard errors by including valid control variables which reduce estimation error on the outcome (see Imbens and Wooldridge (2009)).

Combining DID and matching on lagged outcomes

As discussed above, conditioning on past outcomes exploits all information available in panel data. So why should we not follow this approach to improve estimates within a DID framework? For example, Imbens and Wooldridge (2009) argue that lagged regressors could be correlated with the error term, leading to biased estimates as a violation against exogeneity of control variables. For this reason, Imbens and Wooldridge find a matching approach more attractive than DID for panel data. Furthermore, they add that while matching relies on the CIA (or unconfoundedness), DID relies on the common trend assumption while restricting selection on unobservables on time-fixed factors only – two assumptions that they deem irreconcilable.

In contrast, Chabé-Ferret (2017) meticulously investigates the bias from conditioning on pre-treatment outcomes. He suggests using symmetric pre-treatment and post-treatment periods with a preceding step for matching on all pre-treatment periods, emphasizing that past outcomes should be used only if at least three observations of pre-treatment outcomes are available. He explicitly advises against using a single pre-treatment period's outcome value for matching, in line with similar statements in Wooldridge (2012) and Lechner (2011). We follow a suggestion by Wooldridge (2012), who proposes using a function of multiple pre-treatment outcomes. Given that matching on many pre-treatment periods dramatically reduces the set of available matches within the control group for each treated investor, we calculate simple arithmetic means of pre-treatment outcomes in order to reduce matching dimensionality and improve overlap. Furthermore, taking averages might reduce the bias from matching on past outcomes, which could reinforce selection bias, especially when matching is influenced by transitory shocks.

Finally, Chabé-Ferret (2015) and Chabé-Ferret (2017) find that it is crucial to apply DID symmetrically in order to reduce estimation bias with and without matching. Accordingly, we estimate treatment effects on the first treatment month only and include six months for both the pre-treatment and post-treatment periods.

Model 2: DID and Matching

We employ a conservative two-stage matching process that best identifies valid and balanced treatment and control groups. The first stage generates strata of pre-treatment demographic and account characteristics $\boldsymbol{X}_i$, including pre-treatment means of outcome variables $\boldsymbol{Z}_i$. The second stage implements a one-to-one nearest neighbor propensity score matching conducted separately on each stratum. This results in a one-to-one matching of users of the simulation tool and non-users.

The following DID regression relies on a model equal to (18):

$$Y_{itp} = \tau_{M+DID} W_{itp} + \sum_{i=1}^{I} \beta_i \cdot 1_{I_i=i} \tag{20}$$

$$+ \sum_{k=1}^{T} \lambda_k \cdot 1_{k=t} + \sum_{\ell=1}^{6} \lambda_\ell \cdot 1_{\ell=p} + \boldsymbol{X}'_{itp}\boldsymbol{\delta} + e_{itp} \tag{21}$$

The important difference is that we cannot (only) use time dummies on actual months, but must generate six pre-treatment and six post-treatment period (p) indicators. Each matching control unit receives the same pre-treatment and post-treatment indicator values ($\cdot 1_{\ell=p}$), inherited from its treated counterpart.

Note that for both regression models FE, equation (18), and Matching+DID, equation (21), averages of pre-treatment are differenced out via investor fixed effects and thus do not enter the regression as part of the investor characteristics in $\boldsymbol{X}_{it}$.

Matching on lagged outcomes

As mentioned before, factors that influence the decision to use the portfolio simulator, i.e. participating in the treatment, are potentially unobserved. Since past outcomes are influenced by time fixed factors, a constant, unobserved as well as observed factors, investors with similar past outcomes are also more likely to be similar in terms of unobserved factors conditioning on observable characteristics. O'Neill et al. (2016) use this reasoning to motivate a lagged dependent variable model (LDV) that simply adds the set of available lags on the outcome variable to the right-hand side of the model in equation (4). Interestingly, they find that the LDV model produces the most efficient and least biased estimates compared to a synthetic control and DID approach if the DID parallel trends assumption does not hold.

Model 3: Sequential Matching on Past Outcomes

We combine the merits of matching with the intuitive logic of adding lagged outcome variables to our regression. This is an approach suitable in the presence of panel data and selection on observables fostered by Wooldridge (2012), who proposes the model next to our first fixed effects model.

Sequential matching relies on assuming unconfoundedness conditional on
past outcomes $\{Y_{i,t-1}, ..., Y_{i,1}\}$ and
past treatments $\{W_{i,t-1}, ..., W_{i,1}\}$.
In other words, the selection into the treatment in period t is exogenous with respect to the outcome in period t if we control for observable investor characteristics, $\boldsymbol{X}_i^t$, which include lags for the outcome, treatments and control variables:

$$D\left[W_{it}|Y_{it}(T_i = 0), Y_{it}(W_i = 1), \boldsymbol{X}_i^t\right] = D(W_{it}|\boldsymbol{X}_i^t) \tag{22}$$

Wooldridge proposes estimating the treatment effect for each period on each cross sectional data slice by matching.[13] Following Wooldridge's example, we include contemporaneous values and one lag for the treatment indicator and control variables as well as one lag of the current outcome variable and the lagged relative Sharpe ratio loss in all specifications. To estimate the overall average treatment effect, we simply take the arithmetic mean of the period-specific effects $\{\hat{\tau}_{ate,t}|t = 1, ..., T\}$:

$$\hat{\tau}_{ate} = \frac{1}{T}\sum_{t=1}^{T}\hat{\tau}_{ate,t} \tag{23}$$

Finally, appropriate standard errors are estimated using a panel bootstrapping method by resampling cross-sectional units.[14]

3.2. Treatment effect estimation results

Treatment effects using the fixed effects model

Table 5 reports the treatment effect estimation results of model 1, the fixed effect regression discussed in section 3.1. Outcome variables, i.e. the results on different model specifications with varying dependent variables, are reported in the first four rows, with p-values on the treatment effects in parentheses below. Columns represent sample splits. Column one shows the treatment effect on the overall sample, columns two to five show treatment effect coefficients for regressions on sample splits by pre-treatment quartiles of the relative Sharpe ratio loss distribution across all investors. All estimates are based on regressions which include investor and time-fixed effects and a demeaned mailing indicator. Standard errors are clustered at investor level. The treatment indicator is equal to zero in all periods for Non-Users and equal to zero for Sim-Traders until they have traded a simulated position for the first time. Starting with the first month in which a simulation-position is traded, the indicator turns to one and stays equal to one.

We also tested a specification with monthly indicators equal to one in months when Sim-Traders traded at least one simulated position. If an investor has a fixed optimization goal that she tries

13. We use the user written psmatch2 command in STATA to match one nearest neighbor (Leuven and Sianesi (2003)).
14. For details on the implementation in Stata see Wooldridge (2012) and https://www.stata.com/support/faqs/statistics/bootstrap-with-panel-data/.

to achieve over multiple treatment periods, the room for improvement should decline after each step. Given that the treatment effect is averaged over these multiple treatment periods, when using monthly indicators, we expected the treatment effect to become smaller. Indeed, qualitatively, the results do not change, but the magnitudes of the effects are smaller. The results are available upon request. Further, we estimated specifications that add an indicator variable to each model for the "post" treatment period. The post indicator is equal to zero for each individual investor as long as no trading has occurred during the treatment period. Starting with the first month, in which trading activity of an investor is observed, the indicator turns to one and stays equal to one. With its inclusion, our estimation is a panel data equivalent to classical DID models. The treatment coefficients can be interpreted as the difference between the Sim-Traders' and Non-Users' change in efficiency after the first trading activity during the treatment period. The estimates change only marginally, and are shown in Appendix 1. Our fixed effects model includes only three months after the post indicator or treatment indicator turned to one so as to reduce noise from unrelated events. Both specifications, with and without the post-trading indicator, are qualitatively robust to including all periods.

Before we analyze the estimated treatment effects, we recapture the portfolio efficiency before the introduction of the simulation tool as reported in Table 1, Panel C. In particular, we are interested the ex-ante portfolio efficiency for all simulator users who traded simulated positions (Sim-Traders) in the last set of columns. In the sample split by efficiency quartiles (with Q1 being the most efficient group and Q4 the least efficient) we see that the first quartile has the lowest expected excess return of 5.5% p.a and the third quartile shows the highest average return of 6.5%. Apparently, what distinguishes inefficiency is the large difference in risk, i.e. expected portfolio standard deviation, which ranges from 14.2% (Q1) to 35.8% (Q4). The relative change in expected return between the efficiency quartiles Q1 and Q4 is much lower than the relative change in the expected standard deviation.

Column (1) reports the results on the overall sample, including 706 investors who used the simulator and traded at least one simulated position and 34,067 investors in the control group who did not use the tool but traded actively during the treatment period. The treatment effect on the relative Sharpe ratio loss is highly significant with a p-value of 0.1%. Given the average of 28.6% (see Table 1 , Panel C, column 1) in the relative Sharpe ratio loss across investors, the estimated decrease, i.e. improvement, in portfolio inefficiency by 1.6 percentage points is also economically significant. It is reflected in the estimated increase of 0.7% in the Sharpe ratio, which would increase the average from 30.6% to 31.3%. This is not surprising given that the Sharpe ratio loss is equal to one minus the ratio of portfolio-to-market efficiency, with efficiency measured by the Sharpe ratio. More interestingly, the treatment effect estimates of the expected excess return and expected standard deviation reveal the effects of using the simulator on expected portfolio returns versus risk. Using the information in Table 1, Panel C, we see that the expected excess return's average (IDR) is 6.3% p.a. (5.5) and the standard deviation's average (IDR) is 24.1% (27.3). Hence, the estimated treatment effects reported in column (1) of -0.1 percentage points on the expected return and approximately -1.1 percentage points on the expected portfolio standard deviation are also economically relevant, but admittedly not large. In section 4 we will analyze the optimization behavior of simulator users concerning the trade-off between risk and return more thoroughly.

Now we analyze differences in treatment effects across quartiles in pre-treatment efficiency. In the first and second quartiles, columns (2) and (3), the improvement in the relative Sharpe ratio loss is small and we cannot reject the null hypotheses. Nevertheless, decreases in expected returns are larger than the average effect on the overall sample and significant at the 5% level. Improvements in portfolio risk are comparable in size to the overall effect of -1.1 with -0.8 percentage points and -1.0 percentage points for the first and second efficiency quartiles, respectively, both of which are significant with p-values below 1%. The third and fourth quartile represented in columns (4) and (5) are very interesting. They show highly significant improvements in the relative Sharpe ratios loss of -2.7 and -8.7 percentage points, respectively. For Q4 Sim-Traders this is would reduce the relative loss from 53.6% to 46.9%, a considerably large improvement. It is important to note that the treatment effects for expected returns and portfolio risk change in comparison to the first two quartiles. For both, the third and fourth quartiles, the treatment effects on expected returns are statistically insignificant. Moreover, the sign changes from decreasing expected returns to increasing expected returns between the third and the fourth quartiles. While Q3-investors improve risk about the same magnitude as investors in the first two quartiles by 1.1 percentage points, the effect being narrowly significant at the 10% level, investors in Q4 improve the portfolio standard deviation by 3 percentage points, which is significant at the 5% level. The effect is about three times larger than for investors in any other quartile, which is not entirely surprising given that Sim-Traders show a pre-treatment expected standard deviation of 35.9% in Q4 compared to 17.2% in Q2 (see Table 1, Panel C). Nevertheless, the large improvement in the RSRL of 8.7 percentage points seems to be accomplished not only by reducing risk, but also by improving expected portfolio returns.

In summary, it is very interesting to see that investors across all efficiency quartiles significantly reduce risk. All quartiles show efficiency improvements in the RSRL, even though they are statistically significant only for the third and fourth quartiles. While investors in quartiles Q1 to Q3 seem to give up expected returns in a tradeoff for lower risk, investors in Q4 reduce risk significantly, but from very high levels, while still seeking higher returns.

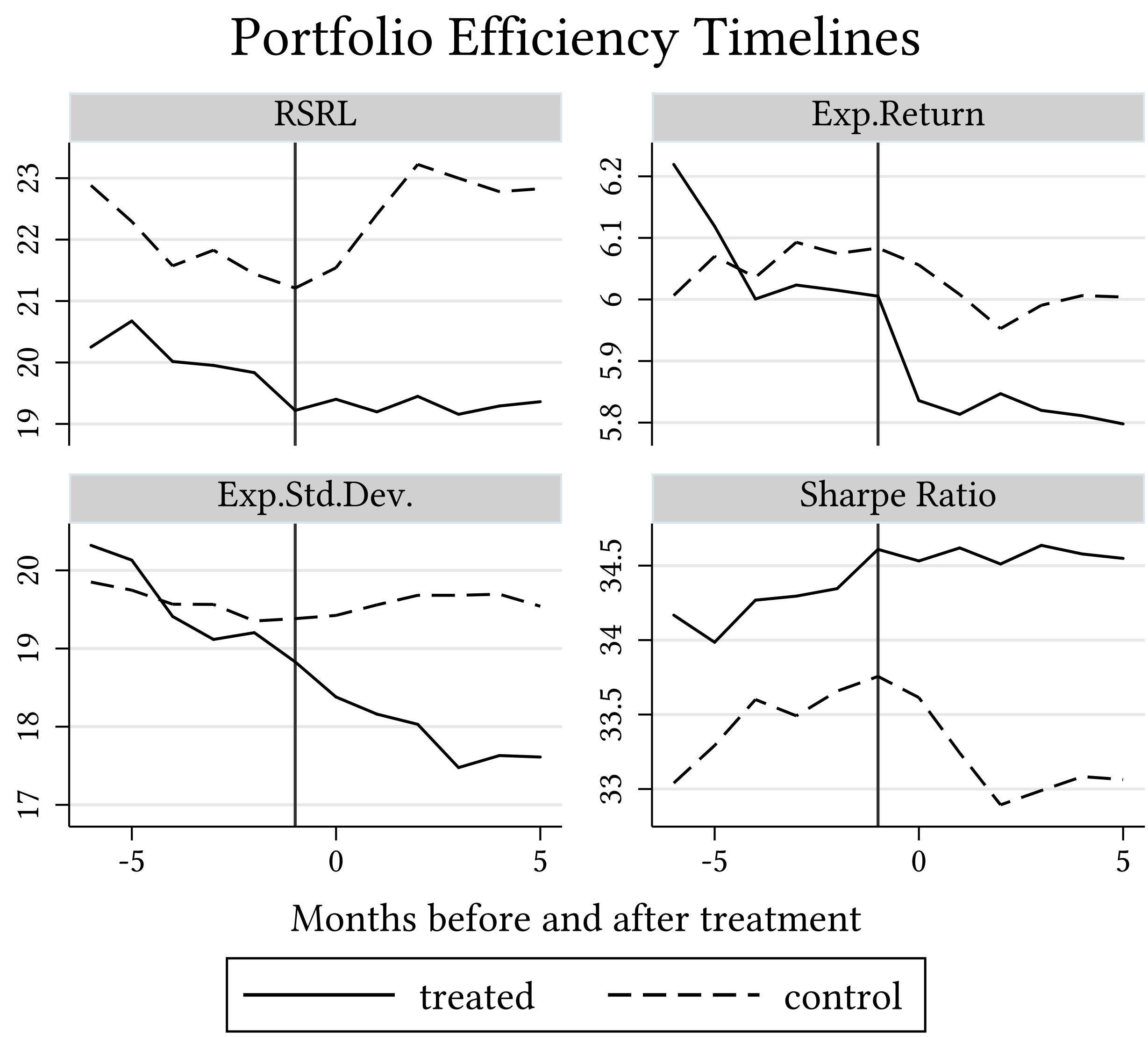
Table 5.
Fixed effects regression: Treatment effects by pre-treatment RSRL quartiles
This table reports the treatment effects estimates for the fixed effect regression model discussed in section 3.1. Outcome variables, i.e. regression models, are sorted in rows together with the corresponding treatment effect coefficients. Column one shows the overall effect, columns two to five show coefficients for regressions on sample splits by pre-treatment quartiles of the relative Sharpe ratio loss distribution across all investors. All estimates are based on regressions that include investor and time-fixed effects and a demeaned mailing indicator. Standard errors are clustered at investor level. The treatment group consists of all investors that traded at least one simulated position (Sim-Traders), the latter group includes all investors who did not log on to any part of the risk management tool but traded at least once after the tool's introduction (Non-Users). The outcome variables are unscaled, for example an excess return of 3.1% enters as 0.031. All coefficients are expressed as a percentage, p-values are provided in parenthesis below. ***, **, and * denote statistical significance at the 1%, 5%, and the 10% levels, respectively.

	(1) All	(2) Q1	(3) Q2	(4) Q3	(5) Q4
Regr. model 1					
Rel. Sharp R. Loss	−1.628***	−0.120	−0.339	−2.653***	−8.659***
	(0.001)	(0.770)	(0.656)	(0.002)	(0.000)
Regr. model 2					
Exp. Excess Return	−0.102**	−0.147**	−0.189**	−0.095	0.173
	(0.026)	(0.012)	(0.039)	(0.314)	(0.263)
Regr. model 3					
Exp. Standard Dev.	−1.067***	−0.796***	−0.994***	−1.078*	−3.001**
	(0.000)	(0.000)	(0.001)	(0.075)	(0.039)
Regr. model 4					
Sharpe Ratio	0.697***	0.052	0.145	1.137***	3.710***
	(0.001)	(0.770)	(0.656)	(0.002)	(0.000)
N-Treated	636	224	177	136	99
N-Controls	34,067	8,119	8,381	8,607	8,960
Controls	✓	✓	✓	✓	✓
Inv FE	✓	✓	✓	✓	✓
Month FE	✓	✓	✓	✓	✓
Clust.SE	investor	investor	investor	investor	investor

Treatment effects using matching

The logit regression on switching between investor groups, Table 4, revealed a wide array of observed investor characteristics which significantly influence the probability of using the simulator and trading simulated positions. To ensure covariate balance and increase the chance of observing parallel trends in the outcome variables between the treated and control groups, we employ a very strict matching regime on fixed investor characteristics and pre-tool introduction averages on time-varying covariates. We match on a dummy for female investors, age, the relationship length, a dummy for Eurex traders, portfolio shares for stocks, equity ETF, bonds, fixed-income ETF, average platform logins, number of securities, log portfolio value, monthly total turnover, and finally, the pre-treatment RSRL quartile, the RSRL value, expected ex-ante return and standard deviation.

In our strict two-stage matching process, we first group similar investors by applying coarsened exact matching (CEM) using the variables mentioned above. Categorical variables create natural groups, while continuous variables are split into quintiles. Within CEM groups we use propensity score matching to find one nearest neighbor for each treated investor. Furthermore, we restrict eligible matches in the control group to investors showing trading activity within the first three months of the post-treatment period of the treated investor. Table A.2 provides summary statistics on the resulting matched sample. The strict matching leaves many treated investors without a match and hence reduces the sample considerably. While the full sample contained 707 SimTraders, matching only supports the inclusion of 342 treated investors. We expected the small sample size to make estimates on the sample splits, especially for Q4 investors who form a group of 48 investors, less stable. The last set of columns in Table A.2 shows the differences and t-test results between the treated and control groups in pre-treatment averages. Differences in most investor characteristics are eliminated by the matching process, guaranteeing covariate balance for the following treatment effect estimation. Notably, in Panel B we see that there is a rather large difference in total portfolio values for Q2 and in turnover for Q4 between the treated and control samples. Given that CEM is performed on quintiles of the corresponding covariate distribution, observations in the tails of the distribution cause these differences. In the case of the portfolio value, this effect is aggravated given that we match on its log transformation. Matching on pre-treatment averages of the outcome variables of interest, namely the relative Sharpe ratio loss, expected portfolio returns and standard deviation, makes it more likely that a common trend is observed, given that we establish similar starting values across treated and controls.



Graphs by Ex-Ante Portfolio Statistic

Figure 4.
Time-trend around treatment
Time trends for the four portfolio efficiency measures on separate y-axes scaled to %. The horizontal axis represents time up to six months before the start of the treatment and five months after the start. The first month of the treatment is at t=0. We use only investors from the matched sample consisting of 342 Sim-Traders (treated) and 342 Non-Users (control). The solid line is the monthly average for the treated group, the dashed line represent the control group.

The common trend assumption is most conveniently assessed by visualizing timelines of the outcome variables for both the treated and control groups. Figure 4 shows time trends for the four portfolio efficiency measures on separate y-axes scaled to %. The horizontal axis represents time up to six periods before the start of the treatment (t=0) and five periods after the start. Given that we use real observational data, the time trends are obviously not perfectly parallel but sufficiently close, especially around t=0. All measures except for the expected return show a promising time trend before the treatment during which portfolio efficiency improves while decreasing portfolio risk. It seems that the treated group continues on this path, while the control group sees a sharp deterioration, for example, in the RSRL after the start of the treatment. Since we match to each treated investor a very similar control group member who traded shortly after the treatment start, it might be the case that unguided trades worsen portfolio efficiency, while assisted trading using the simulation tool improves efficiency. In the Appendix 1, Figure C.1 to Figure C.4, we provide time-trends for all RSRL quartiles. Unfortunately, small sample sizes make time trends very volatile, which indicates that estimates for treatment effects on sample splits should also be interpreted with caution.

Before discussing the regression results for our matching sample, we examine averages on the outcome variables before and after the treatment for both the treatment and the control groups. This enables us to calculate a simple DID treatment effect estimator based on differences (after-before) in differences (Sim-Trader − Non-User). Table 6 reports averages over up to six months before the treatment in the first set of rows and six months after the treatment in the second set of rows. The last set of columns reports group differences and the third set of rows the difference-in-differences. For group differences in the before and after period, standard significance stars for two-sided t-tests at the 10%(*), 5%(**), and 1%(***) levels are provided.

In the first set of rows we observe that statistically significant differences occur for RSRL and Sharpe ratio in the fourth quartile. Even with a very strict matching process which involved a wide array of investor and portfolio characteristics, large differences in outcome variables appear in the before treatment period. The difference of 8.9% is caused by a drop in the RSRL of Sim-Traders in the fourth quartile compared with pre-tool introduction averages. The data reveal that investors with expected portfolio returns close to zero or even negative face large variations in Sharpe ratios and the RSRL in response to portfolio adjustments. Given that we only average over 48 investors in Q4, the impact of individual investors is quite large. While the average standard deviation increased from 21.1% to 25%, the average expected return only increased by 0.4%, which is sufficient to substantially improve the mean RSRL from its pre-tool period average of 50.7% (see Table A.2, Panel B, Sim-Trader Q4) to 41.6% in the before treatment period (see Table 6).

Table 6.
Portfolio efficiency for the matching sample

This table reports averages and differences on the four portfolio efficiency measures. We only include the treatment and control group investors in the matching sample. The first column group for Non-Users (control) shows the summary for all clients with trading activity in the sample period who did not use the simulator or any other component of the risk management tool. The second column group for Sim-Traders (treated) includes clients who have used the simulation tool with at least three simulated portfolios on a single day without trading any simulated position. For each client group we show the overall mean (in column all), and the mean corresponding to the subsample of the RSRL quartile (columns Q1-Q4). The third group of rows represents the differences in means between the treatment and control groups. All values are averaged over six months before (the first four rows) and six months after (rows 5 to 8) the treatment. The difference in means after-before is reported in rows 9 to 12. Test results for t-tests on simple differences are provided as follows: ***, **, and * denote statistical significance at the 1%, 5%, and the 10% levels, respectively. Differences in differences are tested in a regression setup, see Table 7.

	Non-User					Sim-Trader					Sim-Trader - Non-User				
	all	Q1	Q2	Q3	Q4	all	Q1	Q2	Q3	Q4	all	Q1	Q2	Q3	Q4
Before															
Rel. Sharp R. Loss	21.4	10.4	18.0	29.4	50.4	19.7	10.0	18.5	27.0	41.6	−1.7	−0.4	0.5	−2.4	−8.9***
Exp. Excess Return	6.1	5.5	6.3	6.6	6.7	6.0	5.5	6.0	6.6	6.8	−0.1	−0.0	−0.4	0.1	0.2
Exp. Standard Dev.	19.3	14.3	18.2	22.1	32.6	19.0	14.2	17.2	22.0	32.6	−0.3	−0.1	−0.9	−0.1	0.0
Sharpe Ratio	33.7	38.4	35.1	30.3	21.2	34.4	38.6	34.9	31.3	25.0	0.7	0.2	−0.2	1.0	3.8***
After															
Rel. Sharp R. Loss	22.6	12.7	19.4	30.4	48.1	19.3	11.3	19.5	24.8	35.8	−3.3***	−1.4	0.1	−5.6***	−12.3***
Exp. Excess Return	6.0	5.5	6.3	6.2	6.7	5.8	5.2	5.8	6.5	6.8	−0.2	−0.3	−0.5**	0.3	0.1
Exp. Standard Dev.	19.6	15.0	18.7	21.1	33.1	17.9	13.7	16.9	21.0	28.1	−1.7**	−1.3**	−1.8*	−0.0	−5.0
Sharpe Ratio	33.1	37.4	34.5	29.8	22.2	34.6	38.0	34.5	32.2	27.5	1.4***	0.6	−0.0	2.4***	5.3***
After-Before															
Rel. Sharp R. Loss	−1.2	−2.4***	−1.4	−1.0	2.3	0.4	−1.3*	−1.0	2.2	5.7*	−1.6	1.0	0.4	3.2	3.4
Exp. Excess Return	0.1	−0.0	−0.0	0.4	−0.0	0.2	0.2	0.2	0.1	0.1	−0.1	0.2	0.2	−0.2	0.1
Exp. Standard Dev.	−0.3	−0.7	−0.5	1.0	−0.4	1.1	0.5	0.3	1.0	4.5	−1.4	1.2	0.9	−0.1	5.0
Sharpe Ratio	0.5	1.0***	0.6	0.4	−1.0	−0.2	0.6*	0.4	−0.9	−2.5*	0.7	−0.4	−0.2	−1.4	−1.5
N	342	144	87	63	48	342	144	87	63	48	684	288	174	126	96

In the second set of rows, the averages on the after period, Table 6 reports significant differences for the overall matching sample (last set of columns) in the RSRL, standard deviation, and Sharpe ratio. In the after period, the treated group has, on average, improved the portfolio efficiency in comparison to the control group. Finally, below we see that the differences-in-differences show improvements (negative values) on the RSRL over all sample splits. The decrease in the relative Sharpe ratio loss is especially large for the third and fourth quartiles. In the case of Q4 this efficiency improvement seems to be caused by a reduction in portfolio risk, which is estimated to be 5 percentage points. To check the statistical significance of the DID estimates, we report estimation results on regression model 2 in Table 7. Regression estimates closely mirror the simple DID calculations. The overall reduction in the relative Sharpe ratio loss of 1.6%, reported in column (1) is statistically significant at the 5% level, while the decrease of 1.5% in portfolio standard deviation and 0.13% in expected returns are significant at the 1% and 10% levels, respectively. For the first quartile, which includes 144 investors in the treatment and 144 investors in the control group, we report reductions in expected returns and portfolio risk that are highly significant. Given that sample sizes reduce dramatically in the less efficient groups, it is not surprising that effects are not consistently significant. In Q3 and Q4 the treatment effects on the RSRL are larger than 3% but only the improvement for Q3 is statistically significant at the 5% level. For Q4 the large estimated effect of -5% on portfolio risk, i.e. standard deviation, is significant at the 10% level. Comparing these results with the previous fixed effects regression, we obtain qualitatively similar results. Again, the treatment effect on expected returns is worth mentioning. The FE-regression in Table 5 revealed a sign change between the third and fourth quartiles, estimating a positive treatment effect on expected portfolio returns for the fourth quartile that was negative and significant for the first two quartiles. The regression on our matching sample also shows a significant reduction in expected portfolio returns for investors in Q1, but the sign is already reversed for Q3 and Q4 has a small negative treatment effect. Overall, the treatment effect on portfolio efficiency is positive (or negative on the Sharpe ratio loss). We consistently estimate a reduction on the expected portfolio standard deviation. While investors in the most efficient groups in terms of pre-tool RSRL, Q1 and Q2, seem to improve efficiency by trading lower risk against reductions in returns, the mechanism is unclear for investors in Q3 and Q4.

Although using a careful matching process, exogenous time-varying shocks that influence both the outcome and treatment can bias our estimator. In our data, we see that matching on pre-tool averages on the outcome variable, likely does not capture the most recent dynamics that influence treatment and outcomes. The matching variables are not influenced by the treatment and are thus exogenous, but they are 'outdated' if selection depends on time-varying effects. To account for this dynamic environment, we estimate a further robustness check that relies on a sequential matching algorithm. Sequential matching assumes unconfoundedness conditional on past outcomes and past treatments. For each month in the treatment period, a cross-sectional model is estimated by nearest neighbor matching which includes lagged treatment, control and outcome variables. All cross-sectional treatment effects are averaged, and standard errors are estimated by panel bootstrapping. Unfortunately, the method delivered almost no statistically significant effects, resulting from estimated effect sizes which are a fraction of the DID estimates. We provide the results of the sequential matching exercise in Appendix 2, Table B.2. It is not surprising that an approach using monthly indicators delivers smaller effects than a treatment indicator which turns to one and stays at one after the first treatment month. The latter averages outcome values in the pre-treatment and post-treatment periods. Additionally, in previous models, we have only included the first treatment incident, i.e. the first month with traded simulation positions in our fixed effects and matching diff-in-diff regressions. If the effect of using the simulator is getting smaller in a sequence of arbitrary using days, averaging the effects retrieved from monthly indicators would naturally result in smaller estimates. Nevertheless, the effects reported in Table B.2 are very small in comparison to our previous models and hence prone to variations during the bootstrapping resampling which renders almost all estimates statistically insignificant. It is only for Q1 investors that we see a small relative Sharpe ratio loss improvement for the average treatment effect on the treated (ATT), which is barely significant with a p-value of 5.9%. While the average treatment effect estimates show an improvement in the RSRL across investor groups, none of the effects are statistically significant and, given the small effect sizes, even the directions of the treatment effects do not deliver useful information. Given that there is, to the best of our knowledge, no application of the method published so far and that no standard library for statistical computing exists, it is hard to verify our results or assess errors in the estimation and bootstrapping algorithm.

Treatment effects' robustness using a 3-Factor benchmark

Comparing the treatment effects of using the Multi-Asset versus the 3-factor benchmark, we first consider how expected returns and portfolio risk are correlated on monthly portfolio data. Surprisingly, we find that the RSRL has a correlation coefficient of 0.53, while it is 0.58 for expected returns and 0.99 for the portfolio standard deviation. The correlation on our main efficiency measure between the benchmark is thus strikingly low and is driven by differences in the calculated expected returns. We check how investors are sorted into pre-treatment RSRL quartiles and see that of 11,002 investors in the first quartile of the RSRL distribution based on the MultiAsset benchmark, 4,260 appear in Q2, 3,242 in Q3, and 799 in Q4 in the RSRL distribution based on the 3-factor model. It thus makes no sense to generate new sample splits with a different benchmark, and we will rely on the RSRL quartiles based on the multi-asset benchmark.

Strong differences in expected returns between the models do not come as a surprise. The CAPM model calculates security specific betas with the benchmark or risk factors. Betas are equal to the long-term covariance between the security's return and the benchmark return divided by the benchmark's variance. If a security is less correlated with the benchmark, its expected return in the CAPM model is smaller. The 3-factor model consists of the MSCI world and equity risk factors and, hence, by construction will discount the expected returns of non-equity securities. The effect is less strong in the case of the multi-asset benchmark with its 60%-25%-15% asset allocation between equity, fixed income, and commodities. If investors diversify by increasing their fixed income and or commodity share, the treatment effect on portfolio efficiency will more likely be positive and more pronounced using the multi-asset benchmark. Using the 3-factor benchmark, efficiency gains might be counteracted by unfavorable benchmark correlations and thus not recognized at all.

Adding securities to the portfolio that actually have strong diversification benefits, such as fixed-income or commodity ETFs, lead to stronger losses in the expected return when using the 3-factor Benchmark compared with the multi-asset benchmark. This could explain the differences in the RSRL treatment effects between the two benchmarks; compare, for example, the results in Table 5 and Table B.3. To check whether this explanation is plausible, Table 3 shows that both the equity share and the fixed income share is larger in traded simulation positions than in other trades conducted by either Sim-Traders or Non-Users. The direction is thus unclear, but if the increase in equity purchases is driven by the fourth quartile, while fixed income purchases are driven by the first three quartiles, this could explain the observed differences in treatment effects. Without a doubt, analyzing the short-term and long-term trading behavior in conjunction with using the simulation tool would be valuable, but such an exploration would go beyond the scope of this paper.

Summary on the treatment effect estimations

We find that measuring portfolio efficiency based on ex-ante expectations is far from trivial and the choice of benchmark and risk factors heavily influences the inference of our treatment effect estimation. While the treatment effect on the RSRL is not robust across benchmarks, we find that the reduction in risk with accepted lower expected returns is a robust pattern across estimation models. The treatment effect on the expected return for investors in Q4, the worst pre-treatment efficiency group, remains unclear. While we observed insignificant effects with alternating direction in the multi-asset estimations, we can confirm that the estimated effect is greater, i.e. less negative and sometimes positive, for Q4 compared with the other investor groups.

It appears that the treatment effects of using the simulator are different for investors with different optimization schemes and thus preferences. We have seen that investors in Q1 and Q2 already have low portfolio standard deviations before the introduction of the tool. Using the simulator, they might search for securities which match these preferences. Whereas investors holding inefficient, highly risky portfolios might primarily be seeking higher returns and only improve risk because the simulator makes potential efficiency gains salient to them. For both investor types, the simulator might induce a risk reduction, with very different preferences and trading motives. We want to examine these trading motives in the next section.

Table 7.
DID regression with matching on pre treatment averages
This table reports the treatment effect estimates for the DID-matching approach. Matching is based on the time-constant investor characteristics, and pre-treatment averages on portfolio statistics and efficiency measures (see text for details). We include fixed effects for investors, months, and pre-treatment and post-treatment periods. The treatment effect is measured by the interaction of the treated group indicator (Sim-Traders) with a post-treatment indicator. We use a two-step matching algorithm. First, coarsened exact matching (CEM) to ensure covariate balance by stratifying the sample and omitting individuals without a control group member in their stratum. Second, nearest neighbor propensity score matching assigns each treated individual to the most similar control match from its stratum. To control for general improvements in product quality which might induce efficiency gains on any occasion when a client invests her savings, we condition the matching additionally on trading activity within the same month that the treated individual traded a simulated position. To detain complexity, we use only the first instance of trading a simulated position after the tool's introduction in June 2014 for each treated individual. The given month serves as the start of the post-treatment period. We use a maximum of 12 time periods per individual depending on data availability, six pre-treatment and six post-treatment periods. Given that we rely on ex-ante efficiency measures, this is more than sufficient. Each match receives the indicator values from its corresponding treatment group member. Standard errors are clustered at investor level. All other table specifications are identical to Table 5. All coefficients are expressed as a percentage, p-values are provided in parenthesis below. ***, **, and * denote statistical significance at the 1%, 5%, and the 10% levels, respectively.

	(1) All	(2) Q1	(3) Q2	(4) Q3	(5) Q4
Regr. model 1					
Rel. Sharp R. Loss	−1.539**	−0.955	−0.190	−3.085**	−3.618
	(0.015)	(0.205)	(0.876)	(0.023)	(0.131)
Regr. model 2					
Exp. Excess Return	−0.132*	−0.247***	−0.196	0.184	−0.061
	(0.093)	(0.007)	(0.217)	(0.331)	(0.839)
Regr. model 3					
Exp. Standard Dev.	−1.480***	−1.138***	−0.886	−0.166	−5.033*
	(0.003)	(0.001)	(0.134)	(0.852)	(0.081)
Regr. model 4					
Sharpe Ratio	0.659**	0.409	0.081	1.322**	1.550
	(0.015)	(0.205)	(0.876)	(0.023)	(0.131)
N-Treated	342	144	87	63	48
N-Controls	342	144	87	63	48
Controls	✓	✓	✓	✓	✓
Inv FE	✓	✓	✓	✓	✓
Month FE	✓	✓	✓	✓	✓
Clust.SE	investor	investor	investor	investor	investor

4. Do investors actively optimize portfolio efficiency?

In this section, we analyze the simulation tool data which include not only the portfolio holdings entering the simulator as inputs, but also the outcomes as presented in the graphical display of the tool. We primarily focus on these metrics observed directly by investors while using the simulator. As described above, the simulator is a simple two-dimensional graph that visualizes differences between portfolios in terms of risk, measured as the value at risk, and the estimated monthly return. To obtain a comparable efficiency measure, we additionally calculate the standard deviation directly from the value at risk, assuming that returns have zero means. Below, when using the term "the simulator metrics", we are referring to the three simulator output variables: Expected return (E[R]), Value at Risk (VaR), and Standard Deviation (STD).

4.1. How do simulation tool users optimize their portfolios?

We begin the within-analysis of the simulation tool by visualizing the average risk and return of simulated portfolios relative to the actual portfolios that appear as starting points in each simulation session. Figure 5 illustrates the average direction in which simulated portfolios appear after investors have modified portfolio positions within the simulator. For each pre-treatment RSRL quartile within the Sim-Traders and Sim-Users group, the start of the arrow represents the average position of the actual portfolio and the tip of the arrow represents the position of the average simulated position.

Naturally, the range in risk and return grows larger from the first to the fourth quartile, which explains why the arrows are close for Q1 and much more separated for Q3 and Q4. We observe that, on average, both Sim-Traders and Sim-Users increase portfolio risk and expected returns across all RSRL quartiles. Admittedly, Sim-Users in Q3 and Q4 show the smallest change in simulated portfolios relative to the actual starting portfolios. We recall that only Sim-Traders traded at least one simulated position during the treatment-period. It is hence not surprising that Sim-Users did not simulate portfolios with larger changes compared to their actual holdings. Nevertheless, it is striking to see that all groups head in the same direction, apparently searching for higher returns while accepting the risk.

In theory, the arrows could take any direction if investors were completely clueless about the risk and return properties of securities which they add to their portfolio during simulations. Trial and error would eventually show a portfolio position in the risk-return diagram that pleases the simulation user. The average over all trials could also take any direction. It is therefore important to note that in our analyses we only consider simulations that closely resemble the actual portfolios of investors. Possibly, simulation users add single securities separately to the simulator in order to assess their individual risk and return properties before simulating the final aggregated portfolio.

To answer the question of how investors optimize portfolios when using the simulator, we averaged the simulation metrics for series of simulations conducted on a single day which resulted

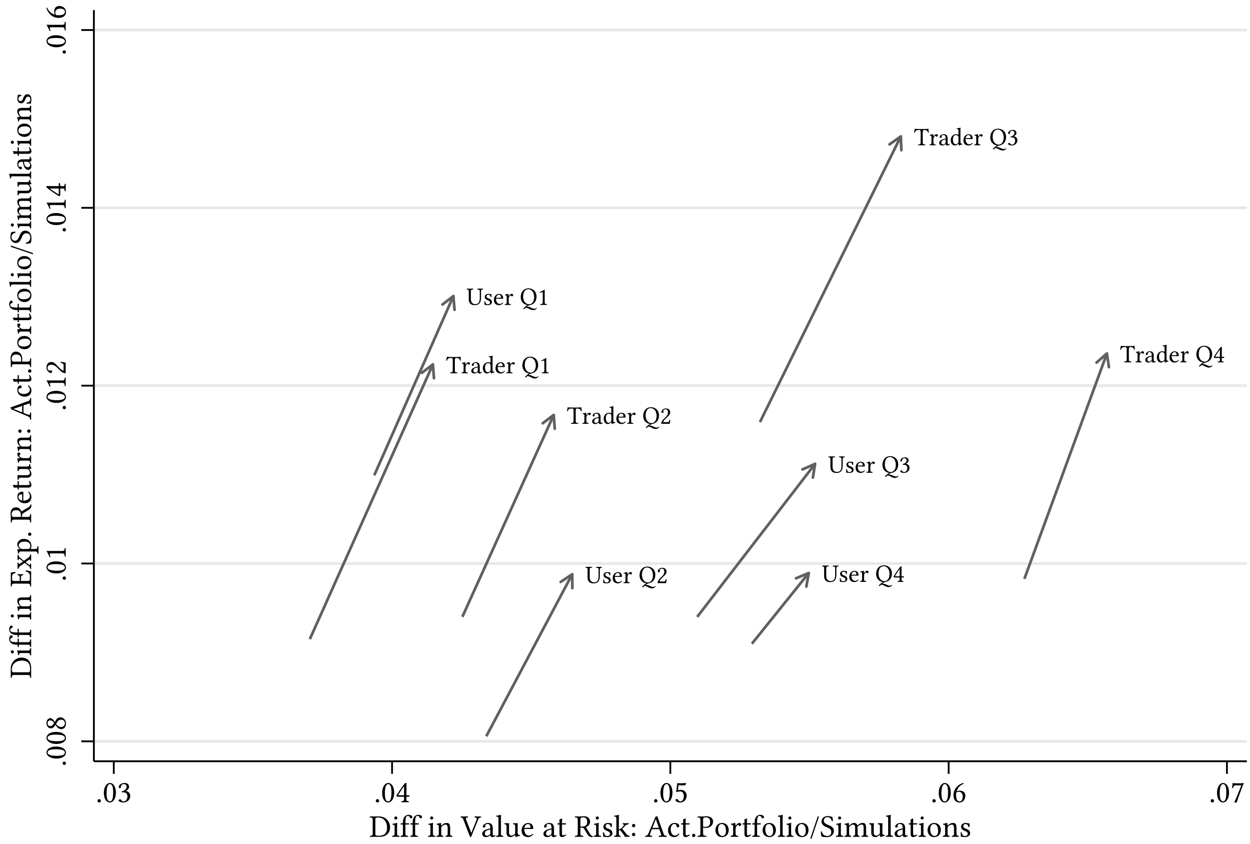


Figure 5.
Actual vs. Simulated Portfolios
The figure illustrates the average direction in which simulated portfolios appear after investors have modified their portfolio holdings within the simulator on first use (for Sim-Users) or trading day (for Sim-Traders). We include only days with at least three and fewer than 50 portfolio simulations. The x-axis represent the difference in the simulator metrics' value at risk and the y-axis the difference in expected return between simulated - actual portfolios. For each pre-treatment RSRL quartile within the Sim-Traders and Sim-Users group, the start of the arrow represents the average position of the actual portfolio and the tip of the arrow represents the position of the average simulated position.

in trading at least one simulated position. Figure 6 illustrates the timelines of these simulation series consisting of the actual starting portfolio, the first, second and third simulations of the day, as well as the simulations that included traded positions. The calculated Sharpe ratio timeline is depicted in the separate, lower panel because of differences in scaling. The timeline trends in simulator metrics, from actual portfolios, over the first three simulations, to the traded simulations, confirm that investors, on average, optimize their starting portfolio towards a portfolio with higher expected return and value at risk. The risk return trade-off leads to an increase in efficiency, i.e. the portfolio Sharpe ratio is improved as well over the series of simulations.

To test whether the results hold for all RSRL quartiles, we estimate differences in the simulator metrics between actual portfolios and simulated portfolios and show the corresponding t-test results in Table 8. Panel A includes only simulated portfolios which did not include traded positions, whereas Panel B includes only simulations with traded positions. We see that even simulation scenarios without trades show significantly higher expected returns across all RSRL quartiles. In return, higher levels of risk are accepted that lead to significant improvements in Sharpe ratio. The picture is very similar for traded simulations. The increase in expected returns of 0.37 percentage points is highly significant over all quartiles (see Panel B, column 1) and is even larger compared for non-traded simulations that improve returns by 0.24 percentage points (see Panel A, column 1). The value at risk increases by 0.54 percentage points (Panel B) compared with 0.42 (Panel A). Overall, the Sharpe ratio increases by 3.46 percentage points, reaching 40.8% for traded portfolios on average.

Judging by the value-at-risk levels in simulations in Panel B (row 6), clients differentiate, in particular, over their risk appetite. The value-at-risk increases from 4.25% to 6.21% from the first to the forth quartile. But only investors in the fourth RSRL quartile increase the portfolio value at risk in simulated portfolios by as much as 1 percentage point from already high levels of 5.26% to 6.21% in traded portfolios. The Sharpe ratio for Q4 investors consequently drops by 4.61 percentage points (Panel B, column 5).

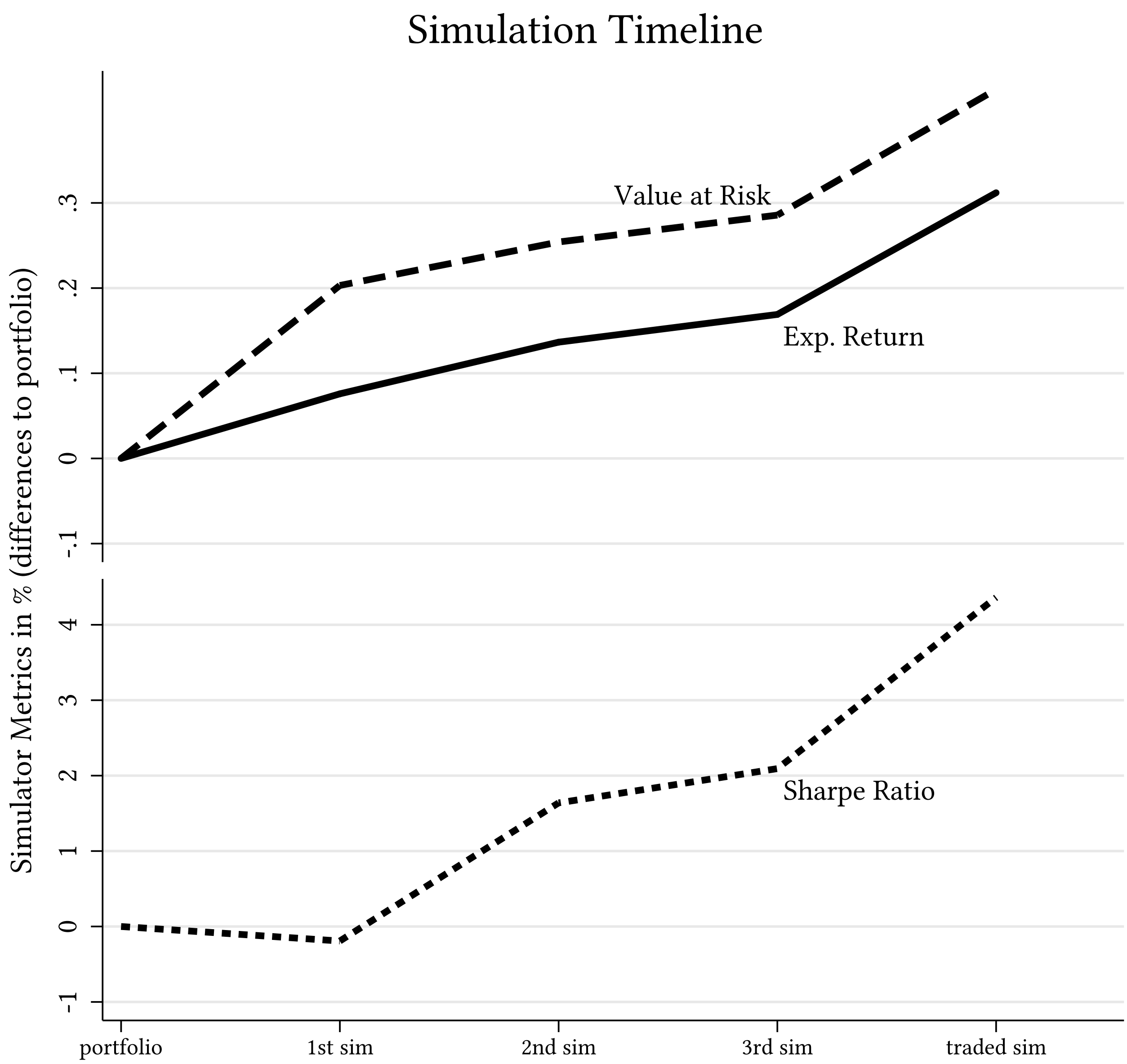


Figure 6.
Simulation Timeline Sim-Traders
The graph shows the difference of the simulator metrics in percentage points between actual portfolios and simulated portfolios. We restrict the sample of simulations to days with at least three and fewer than 50 simulations per Sim-Trader. The first three data points are averages of the corresponding first three simulations per day and the last data point is the average over all traded simulations. The sample consists of 12,282 simulations from 553 investors.

Table 8.
Simulator metrics of simulations and actual portfolios of Sim-Traders without (Panel A) and with traded positions (Panel B):
This table reports the differences in simulator metrics between simulated portfolios and the actual portfolios of Sim-Traders. We restrict simulated portfolios to those closely related to the actual portfolios. Panel A reports means and differences for the simulation that did not contain a traded position. Panel B is restricted to simulations that did contain a traded position. The first column provides the average over all Sim-Traders and the following columns show the averages corresponding to the subsample by the RSRL quartile (columns Q1-Q4). Test results for t-tests on simple differences are provided as follows: ***, **, and * denote statistical significance at the 1%, 5%, and the 10% levels, respectively.

Panel A		All	Q1	Q2	Q3	Q4
Exp. Return	Portf.	1.292	1.210	1.417	1.186	1.480
	Sim.	1.531	1.525	1.593	1.344	1.750
	Sim-Ptf	0.239***	0.315***	0.176***	0.158***	0.270***
Value-at-Risk	Portf.	4.645	3.937	4.883	5.343	5.798
	Sim.	5.070	4.446	5.292	5.594	6.250
	Sim-Ptf	0.425***	0.509***	0.409***	0.251***	0.452***
Sharpe Ratio	Portf.	33.332	37.023	34.276	25.623	28.262
	Sim.	35.599	40.223	35.633	26.798	31.654
	Sim-Ptf	2.266***	3.200***	1.357***	1.175	3.392
N Investors		483	180	133	101	69
N Simulations		16,722	7,045	5,199	3,109	1,369

Panel B		All	Q1	Q2	Q3	Q4
Exp. Return	Portf.	1.381	1.176	1.438	1.722	1.241
	Sim.	1.751	1.550	1.748	2.130	1.697
	Sim-Ptf	0.370***	0.374***	0.310***	0.407***	0.456***
Value-at-Risk	Portf.	4.388	3.772	4.316	5.151	5.255
	Sim.	4.932	4.253	4.691	5.847	6.209
	Sim-Ptf	0.544***	0.481***	0.376***	0.696***	0.954***
Sharpe Ratio	Portf.	37.362	36.886	37.602	36.974	39.253
	Sim.	40.824	41.625	42.238	40.356	34.645
	Sim-Ptf	3.462***	4.738***	4.636***	3.382***	−4.608**
N Investors		519	189	138	108	84
N Simulations		9,189	3,474	2,779	2,002	934

4.2. Do observed simulation tool metrics influence trading?

Our estimates in Table 8 reveal that investors across all efficiency groups seem to optimize their portfolio primarily by searching for higher expected returns. Investors in Q3 and Q4 of the pre-treatment RSRL distribution, are certainly extreme in accepting value at risk levels exceeding 6%. But do investors ignore portfolio risk and focus only on returns? So far, we have only considered differences between simulated and actual portfolios. The actual portfolio is a natural reference point, and an investor with a simulation tool at hand is certainly challenged to optimize the most salient feature, i.e. increase expected returns. The data on conducted simulations do not just reveal the gross direction of optimization, however. We observe both traded and non-traded security positions that were examined in the simulator. We can thus analyze the security positions that simulation users compare and we use these positions to estimate preferences on risk and return. If investors are indifferent to risk and seek only the maximum attainable return given their knowledge about investment options, the portfolio value at risk level of simulated positions should be statistically insignificant. It could be argued that investors probably come across the simulation tool when a trading plan is already fixed and that the simulator is used purely for entertainment and confirmation of ideas. We wish to assess whether investors react to the information provided by the simulation tool by estimating the choice of trading a simulated security position versus not trading it using a simple linear probability model of the following form in order to allow for large dimensional fixed effects:

$$y_{i,s,p} = \boldsymbol{\beta} \boldsymbol{X}_s + \mu_i + \mu_p + \varepsilon_{s,p} \tag{24}$$

For the sake of simplicity, we consider only buy positions. The dependent variable y is equal to one if the security position (p) is traded after simulation (s), and zero otherwise. $\boldsymbol{X}_s$ contains the simulation specific simulator metrics and $\boldsymbol{\beta}$ the preference coefficients of interest. μ_i represent investor x day fixed effects in order to capture trading motives and ideas that are constant over simulations conducted on a single day, but might vary from one day to another. μ_p adds fixed effects for instrument type x month in order to capture market factors such as news or price fluctuations. Standard errors are clustered at investor level. We estimate the model after collecting all security positions added to simulated portfolios that are closely related to the actual portfolios of the investors.

Results are reported in Table 9. Simulator metrics enter the regression as a percentage and coefficients are also reported as a percentage to facilitate interpretation. Column (1) shows the estimates on the overall sample, which reveal that an increase of 1 percentage point in the expected return of the portfolio, including the simulated position, increases the probability of buying the security by 1.37 %, see Model 1. Given that a 1% additional p.a. return on the portfolio is large, the effect size is moderate and is significant only at the 10% level. Likewise, a large improvement of 1% in the value at risk would increase the trading probability by no more than 0.7%; see Model 2.

Interestingly, the coefficients are larger, their size doubles, and they become highly significant if both risk and return enter the model; see Model 4. Our estimates on the sample splits, finally, reveal differences in preferences across the portfolio efficiency quartiles by RSRL. In the rows for Regression Model 4, we see that, with deteriorating portfolio efficiency from RSRL Q1 to Q4, the preference for return increases gradually, while the preference for risk improvements decreases. For Q1 investors a 1% improvement in the value at risk weighs more than a 1% increase in expected returns. From Q3 to Q4 we see the sign on the value at risk coefficient reversed, which would imply ceteris paribus that Q4 investors prefer riskier securities, but the coefficient estimate is not significant. The coefficient on the expected return, on the other hand, is very large at 4.5 and highly statistically significant. For these investors, the trading probability for a security which increases expected portfolio returns by 1%, increases by 4.5%.

In summary, the estimation reveals the investors' preferences over risk and return in simulated portfolios, composed and observed by the group of Sim-Traders. From the previous subsection, comparing original portfolios and simulations, investors across all RSQL quartiles appear to seek higher expected returns only at different levels of risk. In this subsection, comparing simulations with each other, and in particular, comparing traded versus non-traded simulations, we actually see an overall preference against risk (Model 4, column 1 in Table 9). This is driven by clients in the first two quartiles. In contrast, the overall preference across all quartiles for higher returns is confirmed.

Table 9.
Linear Probability Regression of Buying a Simulated Security
The table reports coefficient estimates of linear probability models regressing a dummy for traded simulation positions on observed improvements in the simulator matrix: see Section 4.2, Equation 24. All regressors are scaled to percentages, coefficients are shown as a percentage. Regressors measure improvements in expected return, Value-at-Risk (VaR) and Sharpe ratio, calculated as differences between simulations and the investors' current actual portfolios. For example, an increasing portfolio risk (Value-at-Risk) from 4% (portfolio) to 5% (simulation), is estimated to decrease the probability of trading a simulated security by 0.57% ceteris paribus according to Panel A, model 2, column 1. Results for different model specifications, i.e. simulation metrics entering the regression as independent variables, are displayed in rows. Column 1 shows the overall effect, columns 2 to 5 show coefficients for regressions on sample splits by pre-treatment quartiles of the relative Sharpe ratio loss distribution across all investors. We include daily investor and monthly instrument-type fixed effects. To control for trading motives and market conditions. P-values are provided in parenthesis below. ***, **, and * denote statistical significance at the 1%, 5%, and the 10% levels, respectively.

	(1) All	(2) Q1	(3) Q2	(4) Q3	(5) Q4
Regr. model 1					
Exp. Return	0.943	−0.401	2.480	0.938	4.684**
	(0.194)	(0.738)	(0.147)	(0.412)	(0.017)
Regr. model 2					
Value-at-Risk	−0.573*	−1.022*	−0.949**	−0.408	1.005**
	(0.051)	(0.064)	(0.029)	(0.388)	(0.039)
Regr. model 3					
Sharpe Ratio	0.231***	0.173**	0.330***	0.129	0.333***
	(0.000)	(0.016)	(0.000)	(0.273)	(0.008)
Regr. model 4					
Exp. Return	2.661***	2.006	5.764**	1.893	4.564**
	(0.002)	(0.118)	(0.012)	(0.241)	(0.028)
Value-at-Risk	−1.254***	−1.674***	−2.224***	−0.838	0.081
	(0.000)	(0.005)	(0.000)	(0.221)	(0.862)
N-Inv	553	201	149	116	84
N-Sim.Sec.	62,582	29,039	14,707	13,293	5,532
InvXDay FE	✓	✓	✓	✓	✓
MonthXInstr FE	✓	✓	✓	✓	✓
Clust.SE	investor	investor	investor	investor	investor

Table 10.
Ex Ante Portfolio Efficiency of simulations and actual portfolios of Sim-Traders without (Panel A) and with traded positions (Panel B):
This table reports the differences in simulator metrics between simulated portfolios and the actual portfolios of Sim-Traders. We restrict simulated portfolios to those closely related to the actual portfolios. Panel A reports means and differences for simulations that did not contain a traded position, Panel B is restricted to simulations that did contain a traded position. The first column provides the average over all Sim-Traders and the following columns provide the averages corresponding to the subsample by the RSRL quartile (columns Q1-Q4). Test results for t-tests on simple differences are provided as follows: ***, **, and * denote statistical significance at the 1%, 5%, and the 10% levels, respectively.

Panel A: Not-Traded Simulations		All	Q1	Q2	Q3	Q4
RSRL	Portf.	18.036	10.704	17.771	26.521	37.558
	Sim.	17.456	12.138	17.740	22.134	33.149
	Sim-Ptf	−0.580***	1.434***	−0.032	−4.387***	−4.409***
Exp. Return	Portf.	5.541	5.625	5.068	5.851	6.199
	Sim.	5.493	5.527	5.155	5.859	5.772
	Sim-Ptf	−0.048**	−0.098***	0.086**	0.008	−0.427***
Standard Dev.	Portf.	16.830	14.718	14.401	19.216	31.528
	Sim.	16.751	15.811	14.879	18.139	25.553
	Sim-Ptf	−0.079	1.093***	0.479***	−1.077***	−5.974***
Sharpe Ratio	Portf.	35.117	38.258	35.230	31.481	26.753
	Sim.	35.365	37.643	35.244	33.361	28.642
	Sim-Ptf	0.249***	−0.615***	0.014	1.880***	1.889***
N Investors		483	180	133	101	69
N Simulations		16,713	7,045	5,199	3,100	1,369

Panel B: Traded Simulations		All	Q1	Q2	Q3	Q4
RSRL	Portf.	19.571	11.387	19.639	22.813	42.860
	Sim.	19.080	11.919	20.157	20.979	38.437
	Sim-Ptf	−0.491**	0.532***	0.518*	−1.834***	−4.422***
Exp. Return	Portf.	5.308	5.132	4.495	6.403	6.040
	Sim.	5.357	5.131	4.445	6.434	6.605
	Sim-Ptf	0.049	−0.001	−0.050	0.032	0.565***
Standard Dev.	Portf.	17.578	13.527	13.140	20.030	40.594
	Sim.	17.660	13.658	13.135	19.772	41.480
	Sim-Ptf	0.082	0.132	−0.004	−0.258	0.886
Sharpe Ratio	Portf.	34.459	37.965	34.430	33.070	24.481
	Sim.	34.669	37.737	34.208	33.856	26.376
	Sim-Ptf	0.210**	−0.228***	−0.222*	0.786***	1.895***
N Investors		519	189	138	108	84
N Simulations		9,189	3,474	2,779	2,002	934

4.3. Does the clients' trading on simulator metrics translate into long term efficiency gains?

We recall that we use an ex-ante portfolio efficiency measure with return and risk expectations based on long-term correlations of securities and the multi-asset benchmark in order to estimate objective treatment effects. In contrast, the simulation tool provides simple expectations based on a rolling six-month window of historical returns. With our long-term approach we can objectively measure efficiency gains in the investors' portfolios, but the investors are influenced by the graphical display of the simulator and the metrics provided therein. We compare Table 8, differences between simulations and starting portfolios in terms of observable simulator metrics, and Table 10, differences between simulations and starting portfolios in terms of long-term ex-ante efficiency measures. In particular, we are interested in Panel B, which reports averages for traded simulations.

In summary, only Q3 investors receive long-term efficiency improvements in line with the Sharpe ratio change in terms of simulator metrics. For Q1 and Q2 investors, Sharpe ratio improvements in simulator metrics turn into deteriorations in the long-term RSRL, and for Q4 investors Sharpe ratio deteriorations in the simulator turn into long-term RSRL improvements. Across all client quartiles, investors increase the expected return and risk observed during simulations (see Table 8). Only Q4 clients increase risk to the extent that the calculated Sharpe ratio deteriorates by 4.6 percentage points. Clients in Q1-Q3 increase Sharpe ratios between 3.4 and 4.4 percentage points. If the main goal of clients across all quartiles is to increase expected returns, only Q4 investors successfully translated their optimization strategy from the simulator into long-run success. Their long-term excess return increased on average by 0.565 percentage points, compared to their actual portfolio and significant at the 1% level (see Table 10 column 5). Given the positive but insignificant change in portfolio standard deviation the increase in excess return is certainly driving the significant improvement of 4.4 percentage points in the relative Sharpe ratio loss for Q4 investors.

Surprisingly, investors holding the most efficient portfolios in Q1 and Q2 actually worsen their objective portfolio efficiency. They increase the RSRL by around 0.5 percentage points in traded simulations compared to their starting portfolios. The increase in relative Sharpe ratio loss is economically small, but it is noteworthy that the risk-seeking investors in Q4 actually manage to translate their trading choices into long-term efficiency gains, while Q1 and Q2 fail to do so. Still, the RSRL remains fairly small for clients in Q1, Q2, and Q3, which are 12, 20, and 21, respectively, compared to 38 for Q4 investors. The analysis in section 4.1 shows that investors across all RSRL-quartiles search for higher expected returns and all investors hazard the consequences of higher risk. Notably, the average value-at-risk per quartile in initial starting portfolios as well as traded simulations increase steadily from Q1 to Q4 (see Table 10). The separation of investors into RSRL quartiles is evidently a consequence of the heterogeneous willingness to take risk.

How can we explain the dispersion between short-term and long-term efficiency? Heterogeneity and hence potential lack of trust in the simulator across investors could be considered as a potential barrier to benefiting from using the tool.

We find that the average number of total simulations and simulations turned into actual trades per investor is smaller for Q4 investors, 24 and 10, respectively, compared with investors in other quartiles. Q1 investors, for example, simulated 49 portfolios on average and turned 16 into trades; see Table 11, Panel A, rows 1-4. This indeed indicates differences in the likelihood of using and relying on the simulator depending on the pre-tool introduction portfolio efficiency of investors, i.e. the RSRL quartiles. We further investigate the success of turning short-term improvements in the simulator into long-term improvements, and observe that only around 50% of improvements in expected portfolio returns convert from short-term to long-term; see Table 11, Panel A row 7. Furthermore, the conversion rate does not improve if we consider only simulations with the traded position; see Panel C, row 6. And improvements in portfolio risk show conversion rates which are only around 10% higher.

We illustrate in Figure 7 how differences between simulations and portfolios in terms of short-term and long-term portfolio metrics relate to each other. Each dot represents a simulated portfolio. And each investor contributes at most 2 datapoints: All simulations but the first simulation per investor without traded positions and the first simulation with traded positions are excluded in order to avoid overcrowding the scatter plot. The x-axes represent short-term changes, i.e. the simulator metrics observed by the investors. The y-axes represent long-term changes, i.e. objective, ex-ante measures. The black line represents linear estimates of regressing long-term and short-term simulations to portfolio deltas.

The scatter plot shows that the relationship between short-term and long-term portfolio metrics is nearly random. Even if the correlation of long-term and short-term success depends on the security type added during simulations, the general disconnectedness of both metrics is a dominant explanation for why investors do not convert their portfolio optimization efforts into long-term efficiency gains.

Figure 7.
Simulation delta in long-term vs short-term risk and return
The figure shows long-term vs short-term changes in portfolio metrics for Sim-Traders' simulated portfolios. We calculate changes in portfolio efficiency metrics of simulations with respect to the investors' actual portfolios. The x-axes represents short-term changes, i.e. the simulator metrics observed by the investors. The y-axes represent long-term changes, i.e. objective, ex-ante measures. The plotting area is restricted in order to properly display the bulk of data points, at the expense of outliers. We restrict simulated portfolios to those closely related to the actual portfolios and drop all simulations but the first simulation per investor without and with traded positions, respectively, in order to avoid overcrowding the scatter plot. The black line represents linear estimates of regressing long-term and short-term simulation to portfolio deltas.

Table 11.
Simultaneous short-term and long-term efficiency gains in simulations
Overall (Panel A), without traded positions (Panel B), with traded positions (Panel C):
This table reports the number of simulations overall (column 1) and across RSRL quartiles (columns 2-5) that achieved improvements in terms of short-term, i.e. simulator metrics, expected returns (second set of rows) and risk (third set of rows) and the corresponding percentage of simulations that simultaneously achieved improvements in long-term, i.e. long-term ex-ante, expected returns and risk over the starting portfolio. We restrict simulated portfolios to those closely related to the actual portfolios. Panel A provides numbers for all simulations for which both simulator metrics were available and ex-ante returns could be calculated. Panel B and C are restricted to the first simulation per investor without and with traded positions, respectively, in order to avoid a bias stemming from very active users with many simulation runs.

Panel A: All Simulations	All	Q1	Q2	Q3	Q4
N Investors	595	213	163	126	93
N Simulations	25 755	10 481	7869	5102	2303
Non-traded	16 612	7012	5131	3100	1369
Traded	9143	3469	2738	2002	934
Return short-term improvement	20 747	8838	6140	3886	1883
short- AND long-term improvm.	10 916	4346	3699	1930	941
long/short (%)	53	49	60	50	50
Risk short-term improvement	5872	2360	1705	1302	505
short- AND long-term improvm.	3707	1551	1051	759	346
long/short (%)	63	66	62	58	69

Panel B: Non-Traded Simulations	All	Q1	Q2	Q3	Q4
N Investors	481	179	132	101	69
N Simulations	481	179	132	101	69
Return short-term improvement	313	115	89	67	42
short- AND long-term improvm.	160	58	53	32	17
long/short (%)	51	50	60	48	40
Risk short-term improvement	138	56	32	36	14
short- AND long-term improvm.	95	41	19	25	10
long/short (%)	69	73	59	69	71

Panel C: Traded Simulations	All	Q1	Q2	Q3	Q4
N Investors	517	188	137	108	84
N Simulations	517	188	137	108	84
Return short-term improvement	399	158	100	78	63
short- AND long-term improvm.	212	81	49	50	32
long/short (%)	53	51	49	64	51
Risk short-term improvement	132	43	40	28	21
short- AND long-term improvm.	78	25	24	15	14
long/short (%)	59	58	60	54	67

5. Conclusion

Our estimated treatment effects for using the simulator suggest improvements in long-term portfolio efficiency measured in terms of the Relative Sharpe Ratio Loss (RSRL). But even our careful identification strategy using a difference-in-differences approach and matching on observables, cannot rule out a selection bias. Therefore, we concentrate on analyzing investors' behavior when using the simulator and the resulting simulated portfolios. We analyze, 1. the differences between the actual and simulated portfolios that are visible to users of the simulation tool, i.e. differences in the simulator metrics (see Table 8), 2. the trading choice between simulated alternatives and the resulting revealed preferences for risk and return (see Table 9), and 3. the objective change in efficiency of traded simulations (see Table 10).

First, simple differences in simulations and portfolios show that investors, across all pre-treatment efficiency groups, simulate security positions which, on average, increase both risk and return, at least provided the information in the simulator. Second, analyzing the choice set and trading decisions of simulated positions we find strong heterogeneity in preferences across investor groups. We observe that investors primarily consider securities that increase returns but also portfolio risk (Table 9). Investors holding the most efficient portfolios have balanced preferences over risk and return, whereas the investors with least efficient portfolios are, in fact, risk-seeking. Third and finally, simulations compared with portfolios by our objective, long-term efficiency measures show a small deterioration in efficiency for Q1 and Q2 investors (see Table 10, Panel B, columns 2 and 3), but efficiency gains in terms of the relative Sharpe ratio loss for the less efficient investors in Q3 and Q4 (columns 4 and 5). Investors are not able to consistently convert their short-term optimization in terms of simulator metrics into long-term efficiency gains because the two sets of metrics are almost completely disconnected.

In conclusion, we find investors to use the sandbox and react to the aggregated portfolio information provided in a simple graphical display in line with their individual preferences, which they reveal during simulations. All investors who use the simulator have in common that they indeed actively optimize their portfolio, starting from very different points in the risk-return space, in the direction of higher returns, while accepting the risk associated with it.

The crucial role of appropriate information provision in form of aggregate portfolio measures is revealed by the investors' reaction to the expected risk and return information displayed by the simulation tool. Improvements in portfolio efficiency that simulation tool users observe in terms of the simulation metrics do not consistently translate into improvements in ex-ante efficiency based on CAPM and a multi-asset benchmark. In fact, the relationship is rather random, rendering the simulation tool useless for long-term portfolio optimization.

We conjecture that investors are able to test and choose alternatives which best match their preferences. In the case of the investors holding the least efficient portfolios, this results in even stronger risk-taking and maximizing of expected returns. Therefore, DSS designers should provide the information that nudges users into maximizing objective, long-term efficiency, i.e. into restoring rationality in consumer finance. It is an open question whether the sandbox is a good learning environment and using it leads to long-term adjustments in the trading behavior of investors. This will certainly be a rewarding topic for future research.

Bibliography

Abadie, Alberto, and Guido W. Imbens. 2006. "Large Sample Properties of Matching Estimators for Average Treatment Effects". *Econometrica* 74 (1): 235–267.

Agnew, Julie R., and Lisa R. Szykman. 2005. "Asset Allocation and Information Overload: The Influence of Information Display, Asset Choice, and Investor Experience". *Journal of Behavioral Finance* 6 (2): 57–70.

Angrist, Joshua D., and Jörn-Steffen Pischke. 2008. *Mostly harmless econometrics: An empiricist's companion.* Princeton university press.

Angrist, Joshua David, and Jörn-Steffen Pischke. 2015. *Mastering 'metrics: The path from cause to effect.* Princeton, NJ: Princeton Univ. Press.

Bailey, Warren, Alok Kumar, and David Ng. 2011. "Behavioral biases of mutual fund investors". *Journal of Financial Economics* 102 (1): 1–27.

Barber, Brad M., and Terrance Odean. 2000. "Trading Is Hazardous to Your Wealth: The Common Stock Investment Performance of Individual Investors". *The Journal of Finance* 55 (2): 773–806.

———. 2001. "Boys Will Be Boys: Gender, Overconfidence, And Common Stock Investment". *The Quarterly Journal of Economics, MIT Press* 116 (1): 261–292.

———. 2008. "All That Glitters: The Effect of Attention and News on the Buying Behavior of Individual and Institutional Investors". *Review of Financial Studies* 21 (2): 785–818.

Bateman, Hazel, Christine Eckert, John Geweke, Jordan Louviere, Stephen Satchell, and Susan Thorp. 2016. "Risk Presentation and Portfolio Choice". *Review of Finance* 20 (1): 201–229.

Benartzi, Shlomo. 2001. "Excessive Extrapolation and the Allocation of 401(k) Accounts to Company Stock". *The Journal of Finance* 56 (5): 1747–1764.

Bertrand, M., E. Duflo, and S. Mullainathan. 2004. "How Much Should We Trust Differences-In-Differences Estimates?" *The Quarterly Journal of Economics* 119 (1): 249–275.

Bhattacharya, Utpal, Andreas Hackethal, Simon Kaesler, Benjamin Loos, and Steffen Meyer. 2012. "Is Unbiased Financial Advice to Retail Investors Sufficient? Answers from a Large Field Study". *Review of Financial Studies* 25 (4): 975–1032.

Bhattacharya, Utpal, Benjamin Loos, Steffen Meyer, and Andreas Hackethal. 2017. "Abusing ETFs". *Review of Finance* 21 (3): 1217–1250.

Bradbury, M. A. S., T. Hens, and S. Zeisberger. 2015. "Improving Investment Decisions with Simulated Experience". *Review of Finance* 19 (3): 1019–1052.

Bradbury, Meike, Thorsten Hens, and Stefan Zeisberger. 2019. "How Persistent are the Effects of Experience Sampling on Investor Behavior?" *Journal of Banking and Finance.*

Brunnermeier, Markus K., Christian Gollier, and Jonathan A. Parker. 2007. "Optimal Beliefs, Asset Prices, and the Preference for Skewed Returns". *American Economic Review* 97 (2): 159–165.

Calvet, Laurent E., John Y. Campbell, and Paolo Sodini. 2007. "Down or Out: Assessing the Welfare Costs of Household Investment Mistakes". *Journal of Political Economy* 115 (5): 707–747.

———. 2009. "Fight Or Flight? Portfolio Rebalancing by Individual Investors *". *Quarterly Journal of Economics* 124 (1): 301–348.

Campbell, John Y. 2016. "Restoring Rational Choice: The Challenge of Consumer Financial Regulation". *American Economic Review* 106 (5): 1–30.

Chabé-Ferret, Sylvain. 2015. "Analysis of the bias of Matching and Difference-in-Difference under alternative earnings and selection processes". *Journal of Econometrics* 185 (1): 110–123.

———. 2017. "Should We Combine Difference In Differences with Conditioning on Pre-Treatment Outcomes?" *TSE Working Papers.*

da Silva Rosa, Raymond, and Robert B. Durand. 2008. "The role of salience in portfolio formation". *Pacific-Basin Finance Journal* 16 (1-2): 78–94.

Deutsche Bundesbank. 2019. "Vermögen und Finanzen privater Haushalte in Deutschland: Ergebnisse der Vermögens-befragung 2017".

Etheber, Thomas, Andreas Hackethal, and Steffen Meyer. 2014. "Trading on Noise: Moving Average Trading Heuristics and Private Investors". *SSRN Electronic Journal.*

French, Kenneth R., and James M. Poterba. 1991. "Investor Diversification and International Equity Markets". *American Economic Review* 81 (2): 222–226.

Gaudecker, Hans-Martin von. 2015. "How Does Household Portfolio Diversification Vary with Financial Literacy and Financial Advice? Internet Appendix". *The Journal of Finance* 70 (2): 489–507.

Glaser, Markus, and Martin Weber. 2007. "Why inexperienced investors do not learn: They do not know their past portfolio performance". *Finance Research Letters* 4 (4): 203–216.

Glenzer, Franca, Helmut Gründl, and Christian Wilde. 2014. “"And lead us not into temptation": Presentation formats and the choice of risky alternatives”. *ICIR Working Paper Series.*

Goetzmann, William N., and Alok Kumar. 2008. “Equity Portfolio Diversification*”. *Review of Finance* 12 (3): 433–463.

Goldstein, Daniel G., Eric J. Johnson, and William F. Sharpe. 2008. “Choosing Outcomes versus Choosing Products: Consumer-Focused Retirement Investment Advice”. *Journal of Consumer Research* 35 (3): 440–456.

Grinblatt, Mark, and Matti Keloharju. 2000. “What Makes Investors Trade?” *SSRN Electronic Journal.*

Hackethal, Andreas, Michael Haliassos, and Tullio Jappelli. 2012. “Financial advisors: A case of babysitters?” *Journal of Banking & Finance* 36 (2): 509–524.

Hirshleifer, David, and Siew Hong Teoh. 2003. “Limited attention, information disclosure, and financial reporting”. *Journal of Accounting and Economics* 36 (1-3): 337–386.

Imbens, Guido W., and Jeffrey M. Wooldridge. 2009. “Recent developments in the econometrics of program evaluation”. *Journal of economic literature* 47 (1): 5–86.

Jacobs, Heiko, Sebastian Müller, and Martin Weber. 2014. “How should individual investors diversify? An empirical evaluation of alternative asset allocation policies”. *Journal of Financial Markets* 19:62–85.

Jegadeesh, Narasimhan, and Sheridan Titman. 1993. “Returns to Buying Winners and Selling Losers: Implications for Stock Market Efficiency”. *The Journal of Finance* 48 (1): 65–91.

Kamenica, Emir, Sendhil Mullainathan, and Richard Thaler. 2011. “Helping Consumers Know Themselves”. *American Economic Review* 101 (3): 417–422.

Kaufmann, Christine, Martin Weber, and Emily Haisley. 2013. “The Role of Experience Sampling and Graphical Displays on One’s Investment Risk Appetite”. *Management Science* 59 (2): 323–340.

Kumar, Alok. 2009. “Who Gambles in the Stock Market?” *The Journal of Finance* 64 (4): 1889–1933.

Lakonishok, Josef, and Seymour Smidt. 1986. “Volume for Winners and Losers: Taxation and Other Motives for Stock Trading”. *The Journal of Finance* 41 (4): 951–974.

Lechner, Michael. 2011. “The Estimation of Causal Effects by Difference-in-Difference Methods”. *Foundations and Trends(R) in Econometrics* 4 (3): 165–224.

Leuven, Edwin, and Barbara Sianesi. 2003. *PSMATCH2: Stata module to perform full Mahalanobis and propensity score matching, common support graphing, and covariate imbalance testing.*

Long, J. Bradford de, Andrei Shleifer, Lawrence H. Summers, and Robert J. Waldmann. 1990. "Positive Feedback Investment Strategies and Destabilizing Rational Speculation". *The Journal of Finance* 45 (2): 379–395.

Looney, Clayton Arlen, and Andrew M. Hardin. 2009. "Decision Support for Retirement Portfolio Management: Overcoming Myopic Loss Aversion via Technology Design". *Management Science* 55 (10): 1688–1703.

Markowitz, Harry. 1952. "Portfolio Selection". *The Journal of Finance* 7 (1): 77–91.

Mitton, Todd, and Keith Vorkink. 2007. "Equilibrium Underdiversification and the Preference for Skewness". *Review of Financial Studies* 20 (4): 1255–1288.

Mullainathan, Sendhil, Markus Noeth, and Antoinette Schoar. 2012. "The Market for Financial Advice: An Audit Study". *NBER Working Paper,* number 17929.

Nolte, Sven, and Judith C. Schneider. 2018. "How price path characteristics shape investment behavior". *Journal of Economic Behavior & Organization* 154:33–59.

O'Neill, Stephen, Noémi Kreif, Richard Grieve, Matthew Sutton, and Jasjeet S. Sekhon. 2016. "Estimating causal effects: Considering three alternatives to difference-in-differences estimation". *Health services & outcomes research methodology* 16:1–21.

Polkovnichenko, Valery. 2005. "Household Portfolio Diversification: A Case for Rank-Dependent Preferences". *Review of Financial Studies* 18 (4): 1467–1502.

Seasholes, Mark S., and Guojun Wu. 2007. "Predictable behavior, profits, and attention". *Journal of Empirical Finance* 14 (5): 590–610.

Seru, Amit, Tyler Shumway, and Noah Stoffman. 2010. "Learning by Trading". *Review of Financial Studies* 23 (2): 705–739.

Shefrin, Hersh, and Meir Statman. 1985. "The Disposition to Sell Winners Too Early and Ride Losers Too Long: Theory and Evidence". *The Journal of Finance* 40 (3): 777–790.

Thaler, R. H., A. Tversky, D. Kahneman, and A. Schwartz. 1997. "The Effect of Myopia and Loss Aversion on Risk Taking: An Experimental Test". *The Quarterly Journal of Economics* 112 (2): 647–661.

Thaler, Richard H., and Shlomo Benartzi. 2004. "Save More Tomorrow™: Using Behavioral Economics to Increase Employee Saving". *Journal of Political Economy* 112 (S1): S164–S187.

Tversky, Amos, and Daniel Kahneman. 1985. "The Framing of Decisions and the Psychology of Choice". In *Behavioral Decision Making,* edited by George Wright, 82:25–41. Boston: Springer US.

Wooldridge, Jeffrey M. 2012. "Treatment effect estimation with unconfounded assignment". *FARS Workshop, Chicago.* http://www.ruf.rice.edu/~kr10/slides_fars_2012.pdf.

Appendix A - Additional Summary Statistics

A.I. User group difference t-tests

How are user/trader different?

Table A.1.
Investor descriptives (pre-treatment) - Group Differences

This table reports the inter group differences on investor demographics and characteristics (Panel A), portfolio statistics and portfolio allocations (Panel B), and portfolio efficiency (Panel C) by client groups and quartiles of pre-treatment relative Sharpe ratio loss (RSRL). See Table 1 for corresponding averages. Test results for t-tests on mean differences are provided as follows: ***, **, and * denote statistical significance at the 1%, 5%, and the 10% levels, respectively.

Panel A: Investor demographics

	Sim-User - Non-User					Sim-Trader - Non-User					Sim-Trader - Sim-User				
	all	Q1	Q2	Q3	Q4	all	Q1	Q2	Q3	Q4	all	Q1	Q2	Q3	Q4
Investor characteristics															
Age	-1.7^{***}	-1.5^{***}	-1.3^{**}	-2.0^{***}	-2.5^{***}	-2.2^{***}	-2.1^{***}	-2.1^{**}	-2.2^{**}	-2.5^{**}	-0.5	-0.6	-0.8	-0.2	0.0
Relationship (yrs.)	-0.6^{***}	-0.2^{**}	-0.8^{***}	-1.0^{***}	-0.6^{***}	-0.7^{***}	-0.6^{***}	-0.5^{**}	-1.2^{***}	-0.9^{***}	-0.1	-0.4	0.3	-0.2	-0.4
Risk Tolerance	0.1^{***}	0.3^{***}	0.2^{***}	0.1	-0.2^{**}	0.2^{***}	0.3^{***}	0.3^{***}	0.1	0.1	0.1	0.0	0.1	0.0	0.3^{*}
Investor characteristics (%)															
Female	-0.1^{***}	-0.1^{***}	-0.1^{***}	0.0^{**}	-0.1^{***}	-0.1^{***}	-0.1^{**}	-0.1^{***}	0.0^{*}	0.0	0.0	0.0	0.0	0.0	0.0
Eurex Customer	0.0	0.0	0.0	0.0	0.0	0.0^{***}	0.0^{*}	0.0	0.0	0.0	0.0^{**}	0.0	0.0	0.0	0.0
Self Employed	0.0^{**}	0.0	0.0	0.0^{**}	0.0	0.0^{*}	0.0	0.0	0.0	-0.1^{**}	0.0	0.0	0.0	0.0	-0.1
N 1	2,521	863	682	588	388	707	239	200	151	117	707	239	200	151	117
N 2	34,067	8,119	8,381	8,607	8,960	34,067	8,119	8,381	8,607	8,960	2,521	863	682	588	388

Table A.1.
-Continued

Panel B: Portfolio Statistics and Portfolio Allocations

	Sim-User - Non-User					Sim-Trader - Non-User					Sim-Trader - Sim-User				
	all	Q1	Q2	Q3	Q4	all	Q1	Q2	Q3	Q4	all	Q1	Q2	Q3	Q4
Portfolio Statistics															
Login Days p.m.	4.8^{***}	4.6^{***}	5.4^{***}	5.0^{***}	4.0^{***}	5.8^{***}	5.7^{***}	6.4^{***}	5.3^{***}	5.6^{***}	1.0^{***}	1.1^{**}	1.0^{*}	0.2	1.6^{**}
Portfolio Value (k EUR)	11.0	1.4	12.1	−7.0	−3.8	−2.8	−19.6	−15.1	−2.9	−3.2	-13.7^{*}	−20.9	-27.2^{*}	4.2	0.6
Number of Securities	2.2^{***}	1.8^{***}	1.3^{**}	−0.4	0.3	1.9^{***}	2.2^{*}	0.0	−1.0	0.4	−0.3	0.4	−1.3	−0.6	0.1
Portfolio Statistics (%)															
HHI	-6.8^{***}	-3.2^{***}	-2.7^{***}	-1.4^{*}	-6.8^{***}	-6.2^{***}	-1.9^{**}	−2.0	-3.4^{**}	-5.4^{**}	0.5	1.2	0.7	−2.0	1.4
HHI 100	-6.9^{***}	-0.8^{***}	-1.7^{***}	-2.4^{***}	-8.0^{***}	-6.4^{***}	-0.6^{*}	-1.7^{**}	-2.7^{*}	-5.3^{*}	0.5	0.2	0.0	−0.3	2.6
Buy-Turnover	1.2^{***}	1.2^{***}	2.0^{***}	2.4^{***}	2.2^{**}	4.2^{***}	2.4^{***}	5.1^{***}	5.7^{***}	8.1^{***}	3.0^{***}	1.3^{***}	3.1^{***}	3.3^{**}	5.9^{***}
Sell-Turnover	0.6^{**}	0.8^{***}	1.4^{***}	1.4^{**}	1.0	3.1^{***}	1.4^{***}	4.1^{***}	4.0^{***}	7.3^{***}	2.6^{***}	0.6	2.7^{***}	2.7^{**}	6.2^{***}
Portfolio allocations (%)															
Equity	-2.6^{***}	-4.6^{***}	-4.4^{***}	0.4	−2.1	-2.7^{***}	-4.6^{***}	-4.1^{**}	−0.6	−1.0	−0.1	0.0	0.4	−1.0	1.1
Single Stocks	-8.1^{***}	-2.4^{**}	−2.0	−2.0	-8.1^{***}	-6.3^{***}	−0.7	−1.1	0.2	−4.0	1.9	1.7	0.9	2.2	4.0
Equity ETF	1.6^{***}	1.5^{**}	1.0^{*}	0.9^{***}	0.5	2.5^{***}	4.2^{***}	1.4	0.1	0.2	0.8	2.7^{**}	0.4	−0.8	−0.2
Active Equity Funds	2.5^{***}	-4.2^{***}	-3.5^{***}	0.5	2.8^{***}	0.5	-8.3^{***}	-4.0^{**}	−0.7	1.7	-2.0^{*}	-4.1^{*}	−0.5	−1.2	−1.1
Fixed Income	1.5^{***}	0.9^{**}	1.6^{***}	1.1^{*}	4.2^{***}	0.7	1.4^{*}	0.8	−0.7	1.5	−0.9	0.4	−0.8	−1.8	−2.7
Bonds	−0.1	−0.3	0.7^{*}	−0.1	1.4	−0.4	0.0	−0.5	−0.5	0.8	−0.4	0.3	−1.2	−0.4	−0.5
FixInc ETF	0.5^{***}	0.6^{***}	0.3^{***}	0.1	0.9^{***}	0.5^{***}	0.7^{***}	0.7^{***}	0.2	−0.3	0.0	0.1	0.4	0.1	-1.1^{*}
Active FixInc Funds	1.1^{***}	0.7^{**}	0.5^{*}	1.1^{***}	2.0^{***}	0.6^{*}	0.7	0.6	−0.4	0.9	−0.5	0.0	0.0	-1.5^{*}	−1.1
Balanced Funds	3.3^{***}	4.0^{***}	3.2^{***}	0.9^{*}	1.4^{***}	2.8^{***}	3.0^{***}	2.2^{**}	1.8^{*}	1.4	−0.4	−1.0	−1.0	0.9	0.0
Other	-0.8^{***}	0.2	−0.3	-1.3^{***}	−0.8	−0.1	0.8	0.6	−0.9	−0.4	0.7	0.5	0.9	0.5	0.5
N 1	2,521	863	682	588	388	707	239	200	151	117	707	239	200	151	117
N 2	34,067	8,119	8,381	8,607	8,960	34,067	8,119	8,381	8,607	8,960	2,521	863	682	588	388

Table A.1.
-Continued

Panel C: Portfolio Efficiency

	Sim-User - Non-User					Sim-Trader - Non-User					Sim-Trader - Sim-User				
	all	Q1	Q2	Q3	Q4	all	Q1	Q2	Q3	Q4	all	Q1	Q2	Q3	Q4
Ex-Ante Portfolio Efficiency															
Rel. Sharp R. Loss	-6.0^{***}	-0.5^{***}	-0.3^{***}	-0.4^{**}	-2.7^{***}	-5.6^{***}	-0.2	0.0	-0.4	-2.8^{*}	0.4	0.3	0.3	0.1	-0.1
Exp. Excess Return	-0.5^{***}	-0.4^{***}	-0.5^{***}	-0.3^{***}	-0.7^{***}	-0.4^{***}	-0.4^{***}	-0.4^{***}	-0.3	-0.4	0.1	0.0	0.2	0.0	0.3
Exp. Standard Dev.	-5.0^{***}	-1.1^{***}	-1.7^{***}	-1.0^{***}	-6.6^{***}	-4.4^{***}	-1.0^{***}	-1.1^{***}	-1.1	-5.1	0.5	0.1	0.6	-0.1	1.5
Sharpe Ratio	2.6^{***}	0.2^{***}	0.1^{***}	0.2^{**}	1.1^{***}	2.4^{***}	0.1	0.0	0.2	1.2^{*}	-0.2	-0.1	-0.1	0.0	0.0
3-Factor Model Efficiency															
Rel. Sharp R. Loss	-4.0^{***}	-0.6	-1.5^{*}	0.6	-2.7^{*}	-3.7^{***}	-0.5	-0.2	-1.7	-1.3	0.3	0.1	1.3	-2.3	1.5
Exp. Excess Return	-0.4^{***}	-0.2^{***}	-0.3^{***}	-0.3^{**}	-0.4^{**}	-0.4^{***}	-0.3^{***}	-0.2^{*}	-0.1	-0.3	0.1	0.0	0.0	0.2	0.1
Exp. Standard Dev.	-5.0^{***}	-1.1^{***}	-1.6^{***}	-1.0^{***}	-6.7^{***}	-4.6^{***}	-1.1^{***}	-1.2^{***}	-1.2^{*}	-5.4	0.4	0.0	0.5	-0.3	1.3
Sharpe Ratio	1.6^{***}	0.2	0.6^{*}	-0.2	1.1^{*}	1.5^{***}	0.2	0.1	0.7	0.5	-0.1	0.0	-0.5	0.9	-0.6
N 1	2,521	863	682	588	388	706	239	200	150	117	706	239	200	150	117
N 2	34,067	8,119	8,381	8,607	8,960	34,067	8,119	8,381	8,607	8,960	2,521	863	682	588	388

A.II. Descriptives of the matching sample

Table A.2.
Investor descriptives (pre-treatment) - Group differences for the matched sample

This table reports averages and differences on investor variables on investor demographics and characteristics (Panel A), portfolio statistics and portfolio allocations (Panel B), and portfolio efficiency (Panel C) by client groups and quartiles of pre-treatment relative Sharpe ratio loss (RSRL). In Panel A, all values are as of October 2015, in Panel B and C statistics are based on averages by client on end of month data over the pre-treatment period (06/2013-05/2014). We only include the treatment and control group investors in the matching sample as defined in section 3.2. The first column group for Non-Users (control) shows the summary for all clients with trading activity in the sample period that did not use the simulator or any other component of the risk management tool. The second column group for Sim-Traders (treated) includes clients that have used the simulation tool with at least 3 simulated portfolios on a single day without trading any simulated position. For each client group we show the overall mean (in column all), and the mean corresponding to the sub sample of the RSRL quartile (columns Q1-Q4). The third group of rows represents the differences in means between treatment and control group. Test results for t-tests on mean differences are provided as follows: ***, **, and * denote statistical significance at the 1%, 5%, and the 10% levels, respectively.

Panel A: Investor demographics

	Non-User					Sim-Trader					Sim-Trader - Non-User				
	all	Q1	Q2	Q3	Q4	all	Q1	Q2	Q3	Q4	all	Q1	Q2	Q3	Q4
Investor characteristics															
Age	51.8	50.9	53.4	52.2	50.8	51.4	50.7	53.2	51.7	49.9	−0.4	−0.2	−0.2	−0.6	−0.9
Relationship (yrs.)	10.1	10.5	10.3	9.4	9.3	10.1	10.5	10.3	9.4	9.3	−0.0	−0.0	−0.0	−0.1	−0.0
Risk Tolerance	3.8	3.7	4.0	3.8	3.8	3.9	3.9	3.9	3.9	3.9	0.1	0.2	−0.1	0.1	0.0
Investor characteristics (%)															
Female	8.2	11.8	2.3	7.9	8.3	8.2	11.8	2.3	7.9	8.3	0.0	0.0	0.0	0.0	0.0
Eurex Customer	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
Self Employed	17.0	18.1	20.7	12.7	12.5	15.8	15.3	11.5	22.2	16.7	−1.2	−2.8	−9.2*	9.5	4.2
N	342	144	87	63	48	342	144	87	63	48					

Table A.2.
-Continued

Panel B: Portfolio Statistics and Portfolio Allocations	Non-User					Sim-Trader					Sim-Trader - Non-User				
	all	Q1	Q2	Q3	Q4	all	Q1	Q2	Q3	Q4	all	Q1	Q2	Q3	Q4
Portfolio Statistics															
Login Days p.m.	10.5	10.0	12.0	10.2	9.5	10.9	10.2	12.3	10.9	10.2	0.4	0.2	0.3	0.6	0.8
Portfolio Value (k EUR)	120.1	118.8	208.8	55.7	47.9	89.4	116.5	91.5	61.2	41.0	−30.7*	−2.2	−117.3**	5.4	−6.9
Number of Securities	17.5	20.8	19.6	14.1	8.4	15.6	20.9	15.5	9.5	7.4	−2.0	0.1	−4.1*	−4.6**	−1.0
Portfolio Statistics (%)															
HHI	20.9	14.0	16.2	26.3	43.5	22.1	15.9	20.1	23.7	42.2	1.1	1.9	3.9	−2.7	−1.3
HHI 100	11.8	2.7	7.6	19.0	36.9	11.6	2.8	8.4	18.3	35.1	−0.1	0.1	0.8	−0.8	−1.9
Buy-Turnover	6.0	3.3	7.1	8.0	9.1	7.5	4.0	7.2	9.3	16.3	1.6*	0.7	0.1	1.3	7.2*
Sell-Turnover	5.6	3.0	6.9	7.1	8.8	6.7	3.2	6.4	8.5	15.4	1.1	0.2	−0.5	1.3	6.6*
Portfolio allocations (%)															
Equity	77.2	75.6	80.9	77.5	75.3	78.2	75.0	81.3	84.4	74.4	1.0	−0.6	0.4	6.9	−0.9
Single Stocks	51.8	28.1	59.3	77.1	76.2	51.2	26.5	59.7	77.9	75.4	−0.6	−1.7	0.4	0.8	−0.8
Equity ETF	4.2	8.2	2.0	0.9	0.5	4.7	8.7	2.3	1.4	1.2	0.5	0.5	0.3	0.5	0.7
Active Equity Funds	24.3	40.7	22.4	4.8	4.3	25.6	41.4	23.3	8.7	4.3	1.3	0.8	0.9	3.9	−0.0
Fixed Income	5.4	6.4	5.5	3.7	4.5	5.3	6.5	4.7	1.4	8.2	−0.1	0.1	−0.8	−2.3	3.7
Bonds	2.0	1.5	2.4	1.6	3.5	1.8	1.5	1.6	0.9	4.2	−0.2	0.1	−0.8	−0.7	0.7
FixInc ETF	0.4	0.5	0.7	0.0	0.0	0.7	1.2	0.5	0.0	0.3	0.3	0.8**	−0.2	0.0	0.3
Active FixInc Funds	3.0	4.5	2.4	2.1	0.9	2.8	3.7	2.6	0.5	3.6	−0.2	−0.8	0.2	−1.6	2.7
Balanced Funds	5.9	9.3	4.2	4.5	0.8	7.2	10.0	6.7	5.7	1.7	1.3	0.8	2.5	1.2	0.9
Other	5.2	5.2	4.8	4.8	6.3	4.0	5.2	2.4	2.0	6.3	−1.2	−0.0	−2.5	−2.8*	−0.0
N	342	144	87	63	48	342	144	87	63	48					

Table A.2.
-Continued

Panel C: Portfolio Efficiency

	Non-User					Sim-Trader					Sim-Trader - Non-User				
	all	Q1	Q2	Q3	Q4	all	Q1	Q2	Q3	Q4	all	Q1	Q2	Q3	Q4
Ex-Ante Portfolio Efficiency															
Rel. Sharp R. Loss	21.5	9.4	18.1	30.1	52.4	21.4	9.5	18.8	29.8	50.7	−0.1	0.1	0.7*	−0.3	−1.7
Exp. Excess Return	6.0	5.4	6.2	6.4	6.6	5.9	5.5	6.0	6.5	6.4	−0.0	0.1	−0.2	0.1	−0.2
Exp. Standard Dev.	19.0	14.0	18.0	21.6	32.6	18.8	14.2	17.4	21.6	31.1	−0.3	0.2	−0.6	−0.0	−1.5
Sharpe Ratio	33.7	38.8	35.1	30.0	20.4	33.7	38.8	34.8	30.1	21.1	0.0	−0.1	−0.3*	0.1	0.7
3-Factor Model Efficiency															
Rel. Sharp R. Loss	38.9	35.5	33.4	40.8	56.9	38.0	33.9	34.5	39.5	55.1	−0.9	−1.6	1.2	−1.3	−1.8
Exp. Excess Return	4.5	3.6	4.7	5.2	5.6	4.4	3.7	4.6	5.1	5.5	−0.0	0.1	−0.2	−0.0	−0.1
Exp. Standard Dev.	18.7	13.7	17.7	21.2	32.3	18.4	13.9	17.1	21.4	30.8	−0.3	0.1	−0.6	0.2	−1.6
Sharpe Ratio	24.7	26.1	26.9	23.9	17.4	25.0	26.7	26.4	24.4	18.1	0.3	0.6	−0.5	0.5	0.7
N	342	144	87	63	48	342	144	87	63	48					

Appendix B - Additional Regression Tables

B.I. FE Reg DID

Table B.1.
Fixed effects regression: DID specification
This table reports the treatment effects estimates for the fixed effect regression model with post-treatment indicator as discussed in section 3.1. Outcome variables, i.e. regression models, are sorted in rows together with the corresponding treatment effect coefficients. Reporting is identical to table 5, in contrast, this table shows estimates for an adjusted fixed effects model that additionally includes an indicator variable for the first trading activity during the treatment period. The indicator variable can be interpreted like a post-treatment indicator as in classical DID models. It is equal to zero for each individual investor as long as no trading has occurred during the treatment period. Starting with the first month, in that trading activity of an investor is observed, the indicator turns to one and stays equal to one. With it's inclusion our estimation is panel data equivalent to classical DID models.

	(1) All	(2) Q1	(3) Q2	(4) Q3	(5) Q4
Regr. model 1					
Rel. Sharp R. Loss	-1.617^{***}	-0.184	-0.566	-3.021^{***}	-7.761^{***}
	(0.000)	(0.641)	(0.421)	(0.000)	(0.000)
Post	0.466^{***}	1.741^{***}	1.436^{***}	0.973^{***}	-2.226^{***}
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Regr. model 1					
Exp. Excess Return	-0.077	-0.137^{**}	-0.175^{**}	-0.060	0.195
	(0.103)	(0.014)	(0.050)	(0.494)	(0.253)
Post	-0.034^{***}	-0.063^{***}	-0.053^{***}	-0.067^{***}	0.041^{**}
	(0.000)	(0.000)	(0.000)	(0.000)	(0.048)
Regr. model 1					
Exp. Standard Dev.	-0.949^{***}	-0.698^{***}	-1.006^{***}	-1.269^{**}	-2.443^{*}
	(0.001)	(0.001)	(0.001)	(0.024)	(0.060)
Post	0.307^{***}	0.520^{***}	0.545^{***}	0.815^{***}	-0.692^{***}
	(0.000)	(0.000)	(0.000)	(0.000)	(0.009)
Regr. model 1					
Sharpe Ratio	0.693^{***}	0.079	0.243	1.294^{***}	3.325^{***}
	(0.000)	(0.641)	(0.421)	(0.000)	(0.000)
Post	-0.200^{***}	-0.746^{***}	-0.615^{***}	-0.417^{***}	0.954^{***}
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
N-Treated	707	239	200	151	117
N-Controls	34,067	8,119	8,381	8,607	8,960
Controls	✓	✓	✓	✓	✓
Inv FE	✓	✓	✓	✓	✓
Month FE	✓	✓	✓	✓	✓
Clust.SE	investor	investor	investor	investor	investor

B.II. Sequential Matching Regression Result

Table B.2.
Sequential matching results: Treatment effects by pre-treatment RSRL quartiles
This table reports estimates using a sequential matching approach on average treatment effects on the treated (ATT) in Panel A, and average treatment effects (ATE) in Panel B (see section 3.2 for details). Sequential matching assumes unconfoundedness conditional on past outcomes and past treatments. For each month in the treatment period, a cross sectional model is estimated by one-nearest neighbor propensity score matching that includes lagged treatment, control and outcome variables. We include outcome variable lags and pre-treatment period averages on the four portfolio efficiency measures as well as portfolio value, turnover, stock, bonds, equity ETF, and fixed income ETF allocations. We control for time-invariant demographics age and risk class, as well as pre-treatment average logins and security count. Finally, we include a monthly indicators for trading activity and info-mailings. Cross sectional treatment effects over all periods are averaged and standard errors are estimated by panel bootstrapping, which performs resampling of investors including all their datapoints instead of individual datapoints.

PANEL A: Average treatment effect on the treated

	(1) All	(2) Q1	(3) Q2	(4) Q3	(5) Q4
Rel. Sharp R. Loss	0.0735	−0.2679*	−0.1180	0.1434	−0.2249
	(0.613)	(0.059)	(0.468)	(0.408)	(0.462)
Exp. Excess Return	−0.0276*	−0.0366*	−0.0083	0.0079	−0.0567
	(0.061)	(0.089)	(0.740)	(0.792)	(0.174)
Exp. Standard Dev.	−0.1466	−0.1415	−0.0472	−0.0658	−0.2014
	(0.275)	(0.300)	(0.746)	(0.717)	(0.675)
Sharpe Ratio	−0.0046	0.0064	−0.0448	0.0439	−0.0294
	(0.939)	(0.921)	(0.504)	(0.534)	(0.822)

PANEL B: Average treatment effects

	(1) All	(2) Q1	(3) Q2	(4) Q3	(5) Q4
Rel. Sharp R. Loss	−0.0893	−0.1707	−0.1545	−0.1731	−0.1231
	(0.605)	(0.264)	(0.302)	(0.424)	(0.756)
Exp. Excess Return	0.0257	0.0362	0.0033	0.0443	0.0335
	(0.241)	(0.183)	(0.907)	(0.256)	(0.599)
Exp. Standard Dev.	−0.0028	0.0390	−0.0478	0.1024	−0.3326
	(0.990)	(0.822)	(0.737)	(0.732)	(0.645)
Sharpe Ratio	0.0426	0.0702	0.0365	0.0434	0.1158
	(0.553)	(0.187)	(0.521)	(0.596)	(0.450)

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

B.III. Robustness Checks: Treatment effects estimation using 3-Factor Model portfolio efficiency

Table B.3.
3-Factor Model: Fixed effects regression by pre-treatment Multi Asset RSRL quartiles
This table reports the treatment effects estimates for the fixed effect regression model discussed in section 3.1. Outcome variables, i.e. regression models, are sorted in rows together with the corresponding treatment effect coefficients. Column one shows the overall effect, columns two to five show coefficients for regressions on sample splits by pre-treatment quartiles of the relative Sharpe ratio loss distribution across all investors. All estimates are based on regressions that include investor and time fixed effects and a demeaned mailing indicator. Standard errors are clustered on investor level. The treatment group consists of all investors that traded at least one simulated position (Sim-Traders), the latter group includes all investors that did not log in to any part of the risk management tool but traded at least once after the tool introduction (Non-Users). The outcome variables are unscaled, for example an excess return of 3.1% enters as 0.031. All coefficients are in percent, p-values are provided in parenthesis below. ***, **, and * denote statistical significance at the 1%, 5%, and the 10% levels, respectively.

	(1) All	(2) Q1	(3) Q2	(4) Q3	(5) Q4
Regr. model 1					
Rel. Sharp R. Loss	1.163*	1.969***	2.641**	0.657	−4.003*
	(0.074)	(0.007)	(0.032)	(0.567)	(0.081)
Regr. model 2					
Exp. Excess Return	−0.209***	−0.185***	−0.327***	−0.239*	−0.055
	(0.000)	(0.001)	(0.001)	(0.051)	(0.804)
Regr. model 3					
Exp. Standard Dev.	−0.993***	−0.777***	−0.864***	−1.005*	−2.698**
	(0.000)	(0.000)	(0.006)	(0.084)	(0.035)
Regr. model 4					
Sharpe Ratio	−0.469*	−0.795***	−1.066**	−0.265	1.616*
	(0.074)	(0.007)	(0.032)	(0.567)	(0.081)
N-Treated	706	239	200	150	117
N-Controls	34,066	8,119	8,380	8,607	8,960
Controls	✓	✓	✓	✓	✓
Inv FE	✓	✓	✓	✓	✓
Month FE	✓	✓	✓	✓	✓
Clust.SE	investor	investor	investor	investor	investor

Table B.4.
3-Factor Model: DID regression on the Multi-Asset matching sample
This table reports the treatment effect estimates for the DID-matching approach. Matching is based on the time constant investor characteristics, and pre-treatment averages on portfolio statistics and efficiency measures (see text for details). We include fixed effects for investors, months, and pre- and post-treatment periods. The treatment effect is measured by the interaction of the treated group indicator (simulation trader) with a post-treatment indicator. We use a two-step matching algorithm. First, coarsened exact matching (CEM), to ensure covariate balance by stratifying the sample and omitting individuals without a control group member in their strata. Second, nearest neighbor propensity score matching assigns each treated individual the most similar control match from its stratum. To control for general improvements in product quality that might induce efficiency gains on any occasion that a client invests her savings, we condition the matching additionally on trading activity within the same month that the treated individual traded a simulated position. To detain complexity, we only use the first instance of trading a simulated position after the tool introduction in June 2014 for each treated individual. The given month serves as the start of the post-treatment period, we use a maximum of 12 time periods per individual depending on data availability, 6 pre- and 6 post-treatment periods. Given that we rely on ex-ante efficiency measures this is more than sufficient. Each match receives the indicator values from its corresponding treatment group member. Standard errors are clustered on investor level. All other table specifications are identical to table 5. Above all, the coefficients are in percent, p-values are provided in parenthesis below. ***, **, and * denote statistical significance at the 1%, 5%, and the 10% levels, respectively.

	(1) All	(2) Q1	(3) Q2	(4) Q3	(5) Q4
Regr. model 1					
Rel. Sharp R. Loss	0.671	3.060**	2.229	−3.295	−4.031
	(0.529)	(0.010)	(0.310)	(0.158)	(0.341)
Regr. model 2					
Exp. Excess Return	−0.152	−0.373***	−0.277	0.363	0.080
	(0.151)	(0.000)	(0.129)	(0.223)	(0.866)
Regr. model 3					
Exp. Standard Dev.	−1.518***	−1.230***	−0.880	−0.363	−4.801*
	(0.002)	(0.001)	(0.132)	(0.675)	(0.090)
Regr. model 4					
Sharpe Ratio	−0.271	−1.235**	−0.900	1.330	1.627
	(0.529)	(0.010)	(0.310)	(0.158)	(0.341)
N-Treated	342	144	87	63	48
N-Controls	342	144	87	63	48
Controls	✓	✓	✓	✓	✓
Inv FE	✓	✓	✓	✓	✓
Month FE	✓	✓	✓	✓	✓
Clust.SE	investor	investor	investor	investor	investor

Appendix C - Additional Figures

C.I. Timeline graphs for the matching sample

The following figures provide time-lines of the portfolio efficiency measures by RSRL quartile sample splits for the matching sample (see section 3.2 and figure 4). Portfolio efficiency measures on separate y-Axes scaled to %. The horizontal axis represents time up to 6 months before the start of the treatment and 5 months after the start. The first month of the treatment is at t=0. We use only investors from the matched sample consisting of 342 Sim-Traders (treated) and 342 Non-Users (control). The solid line is the monthly average for the treated group, the dashed line represent the control group.

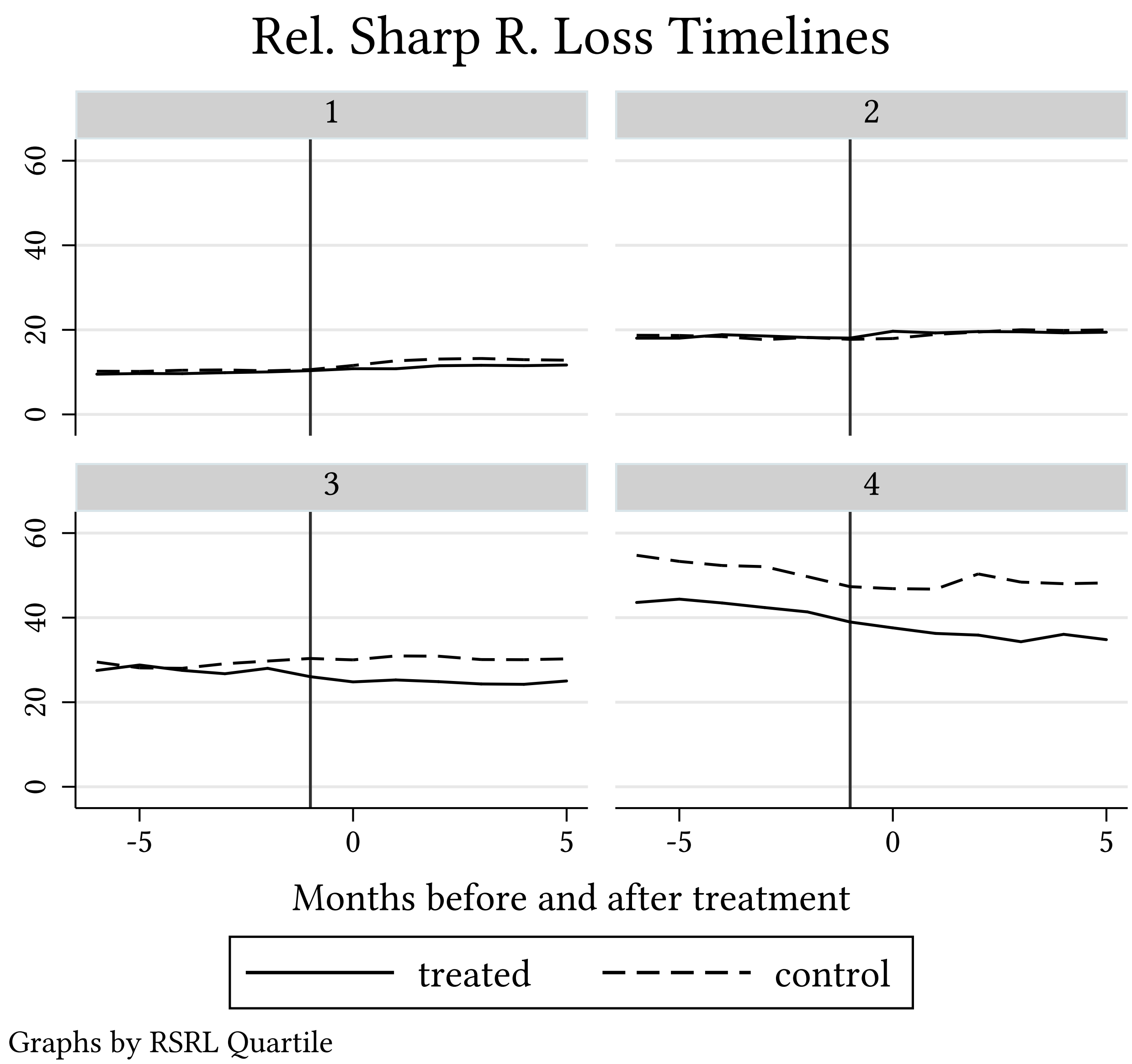


Figure C.1.
RSRL time-trend around treatment

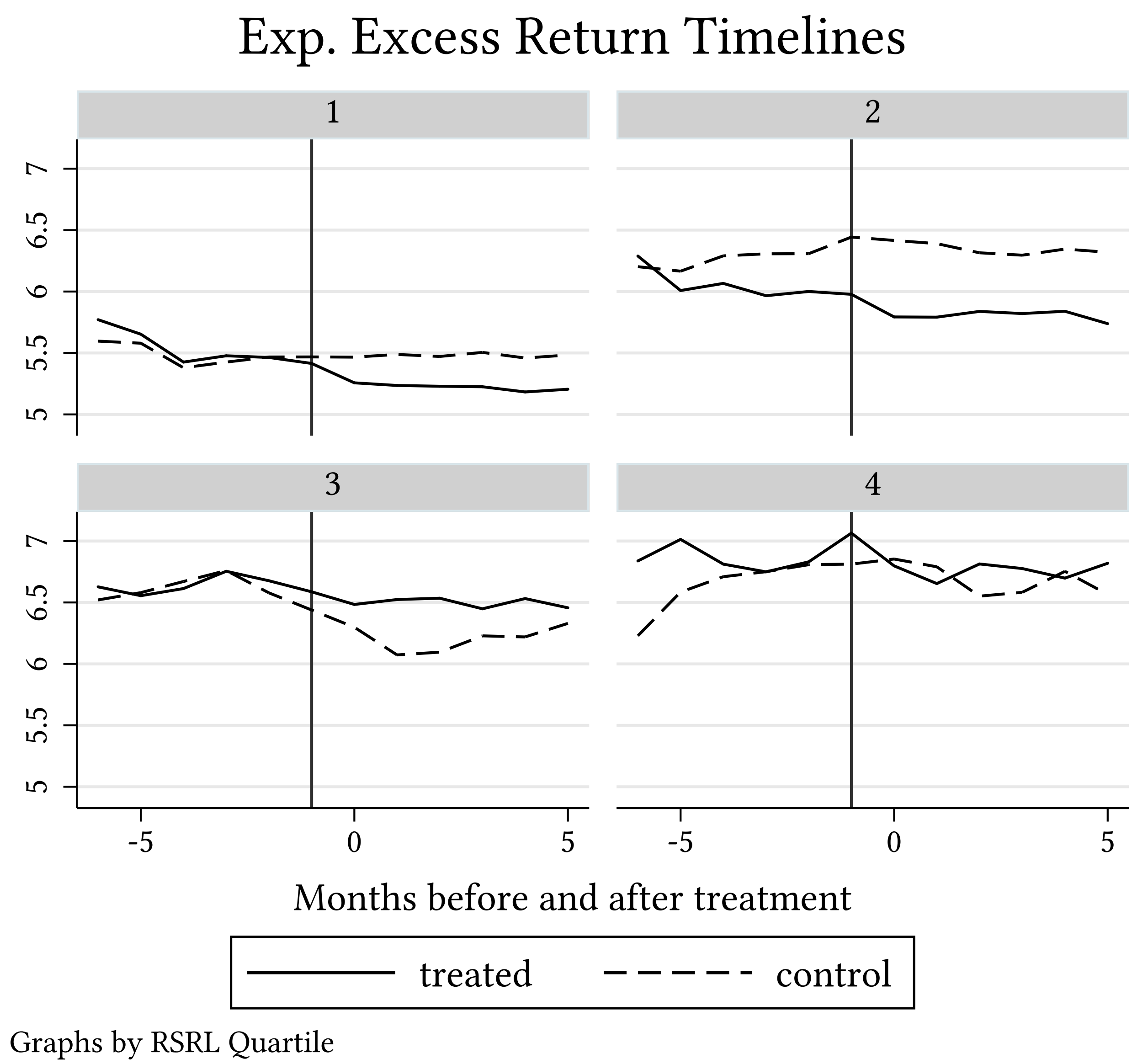


Figure C.2.
Exp. portfolio return time-trend around treatment

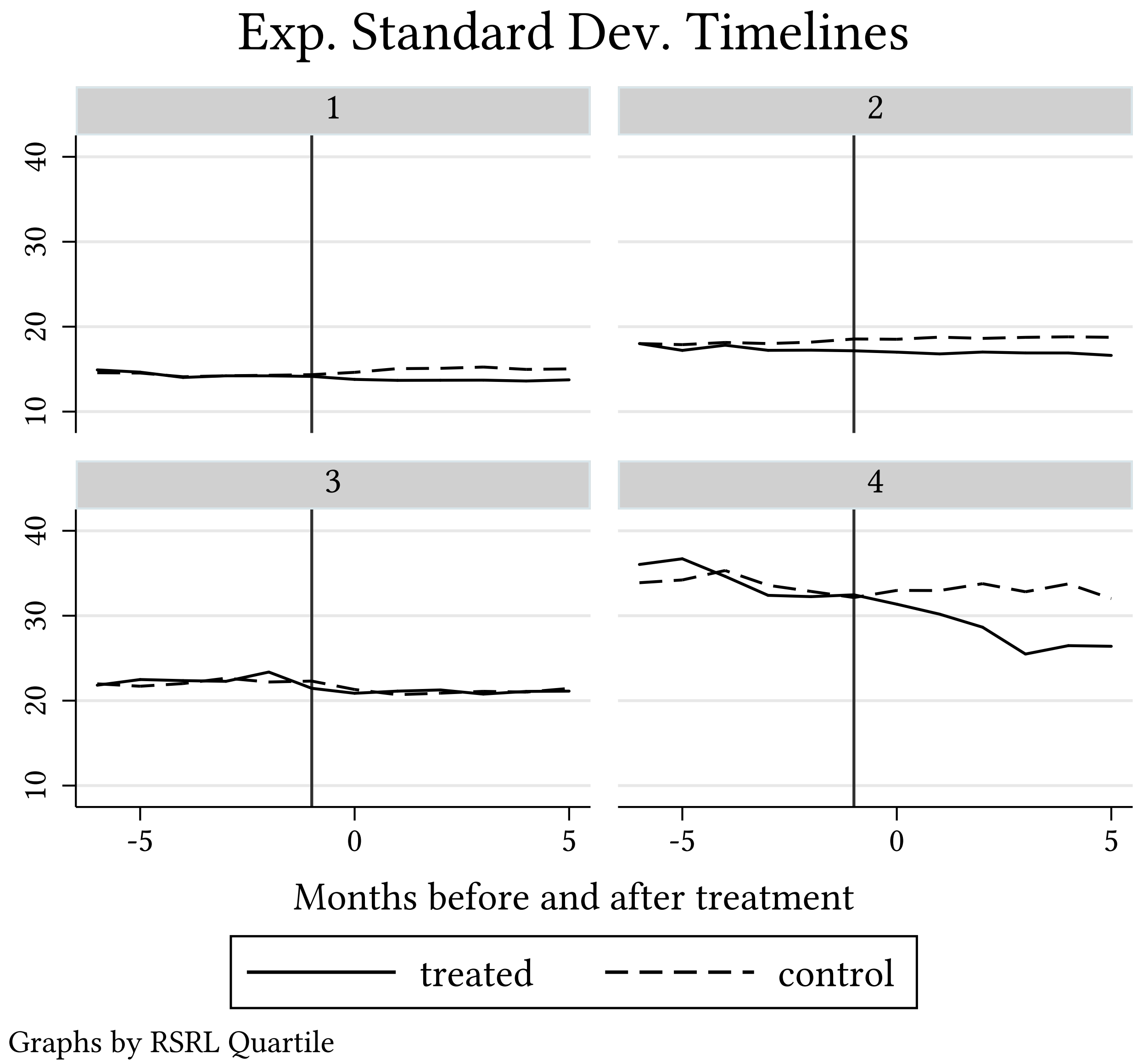


Figure C.3.
Portfolio standard deviation diagram by quartiles of pre-treatment RSRL averages

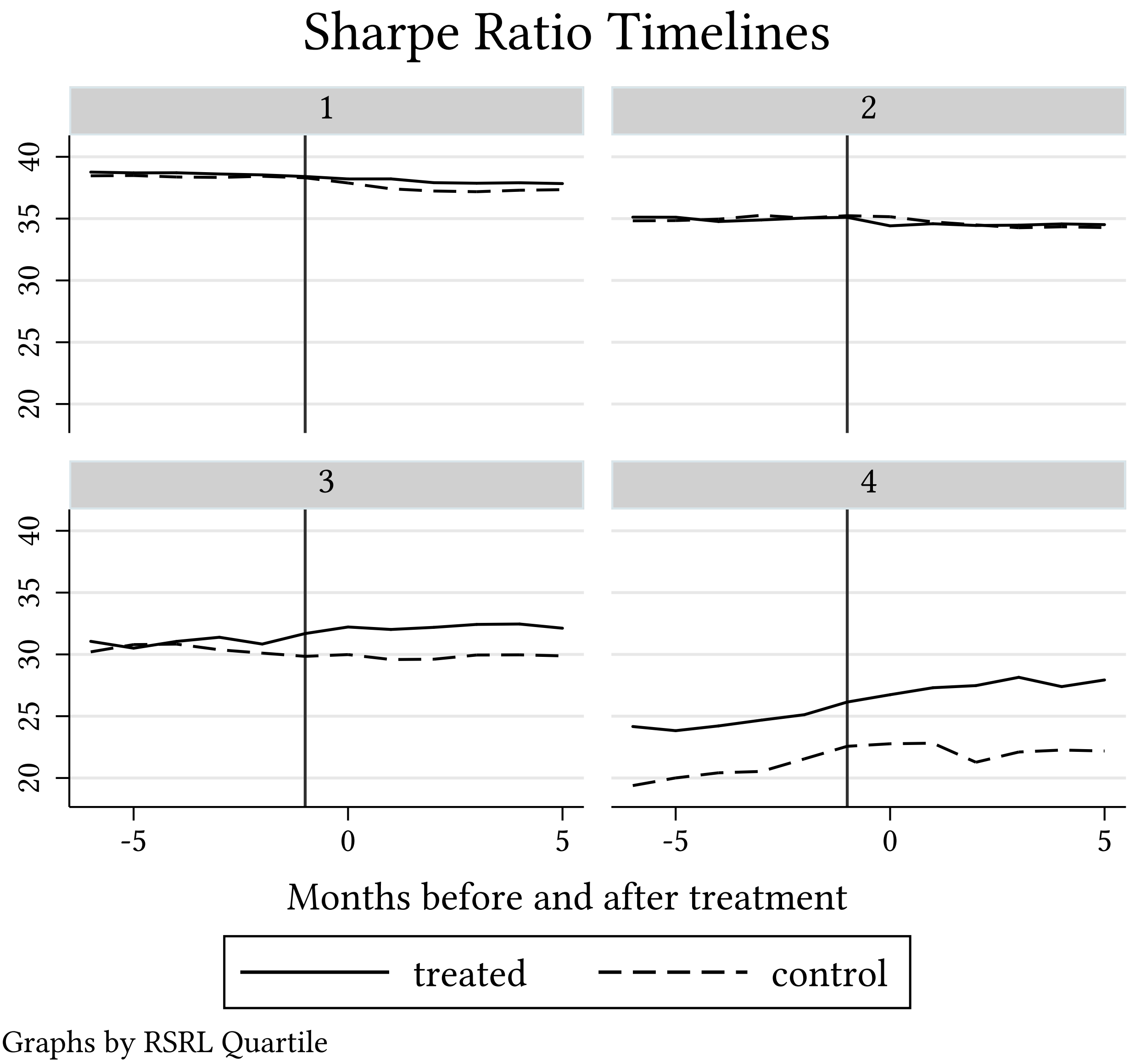


Figure C.4.
Portfolio Sharpe Ratio time-trend around treatment